

Prospect for agreement at Geneva

The latest Soviet proposals on nuclear weapons may let negotiators negotiate. But Britain and France must decide how and when their missiles will be counted. And President Reagan must change his style.

THE most helpful feature of Mr Yuri Andropov's declaration last week on nuclear weapons in Europe is that it is far from being crystal clear. For with the talks on intermediate-range nuclear forces due to resume in Geneva (next week), it is high time that the United States and the Soviet Union abandoned their recent practice of negotiation by means of press releases. The advantage, of course, is that the less clearly a negotiating position is known or understood in advance, the more easily it can be modified when private negotiations begin.

What Mr Andropov is now saying is also, however, a more negotiable proposition than his previous proposal that the numbers of SS20 missiles — Soviet devices carrying up to three warheads each — should be reduced to match the British and French nuclear forces in some way or other. For that line of argument would have required of the Soviet Union's negotiating partner (consisting exclusive of the United States) the complete abandonment of the planned deployment of cruise and Pershing II missiles beginning at the end of this year, the concession (which constitutionally it cannot deliver) that British and French nuclear forces should be regulated by an agreement negotiated only by the superpowers (and France does not even belong to the North Atlantic Treaty Organization (NATO)) and the requirement that NATO itself should eat many of its words, especially the 1979 declaration that West European nuclear forces should be strengthened to match the rapid growth of the SS20 force and the more recent opinion that the French and British nuclear forces are what their governments say they are — strategic. The idea now is that there should be some equality based on counts of warheads, not missiles, between the East and West in Europe, but that the British and French nuclear forces should still figure somewhere in the equation. The proposal is like a fruit cake laced with razor blades, but nevertheless more palatable than pure steel.

Difficulties

The British and French forces are an obvious stumbling block, and are bound to be so. The constitutional difficulty that Mr Andropov brushed aside last week is moreover not a mere formality that might, for example, be resolved by joining the British and French Governments in the negotiations at Geneva. For while the British Government has usually justified its independent nuclear forces as a kind of general contribution to the defence of Western Europe (with luck, by the deterrence of attack), successive governments have implicitly moved closer to the position the French have always taken — that nuclear guarantees by others cannot in all circumstances be relied upon. The Gallois doctrine of the 1950s, that safety in the face of a nuclear threat requires a nuclear retaliatory force commensurate in some sense with its value as a military prize, seems to be gaining more influence than its simplicity deserves. But these independent forces are also technically of a strategic nature. Even though the range of the British missiles carried by four submarines could allow them to be used against targets also accessible to the land-launched missiles due to be introduced to Western Europe in the next few years, their accuracy is of necessity much less.

So the British and French forces belong not in the negotiations on intermediate-range weapons but in the parallel negotiations on strategic arms reductions, called START, also at Geneva. But that is also a bilateral negotiation, while the British Government is

also right to say that its own independent strategic force is so tiny compared with those of the superpowers — a few per cent, according to the basis of counting — that it should be held significant only when plans have been made substantially to reduce the main striking forces. The most that the Soviet Union is likely to win on this point in the intermediate-range negotiations is therefore an undertaking that the independent European forces will be counted along with others at some predetermined point in what it must be hoped will be a continuing process of strategic arms reductions. To its credit, if in slightly lukewarm tones, the British Government has been saying that it will reconsider its own position on nuclear weapons if circumstances should change substantially. (The Government of France has not yet made even such a grudging promise.) Should not both these West European governments now attempt to define the circumstances in which they would acknowledge that their own nuclear forces must be counted explicitly in the START negotiations, not implicitly allowed for as in the numerical imbalance of the Salt II agreement?

Solutions

All this boils down to the assumption that, at some stage, the two sets of negotiations at Geneva must be linked. This has been obvious to outsiders since the two processes began. Now it seems that there can be little progress on the intermediate-range negotiations unless the need for linkage is recognized formally, and the circumstances in which it will occur agreed. But even if a bargain can be struck about the time at which account will be taken of the independent forces, what chance is there that an agreement can be reached on the other problems of intermediate nuclear weapons — verification, the issue of whether the Soviet Union would be required to destroy surplus but mobile SS20s or merely move them out of range of Western Europe, for example? Given that so much of the past few months has been occupied with public statements by the two principals at Geneva, it seems improbable that much can be accomplished before the long diplomatic summer gets under way. But fortunately that is not necessarily the case.

Democratic elections may yet influence the way that events turn out. Mr Andropov's offer to reduce the SS20 force to match the British and French nuclear forces was evidently timed in such a way as to influence the West German elections in April, but seems not to have had that effect. Although the British election on 9 June will be enlivened (or made dull, according to taste) by arguments about unilateral disarmament, there is hardly enough time in which to seek to influence its outcome in any important way. The more influential election is that which must take place in the United States in November 1984, and in which President Reagan may yet be a candidate. In the past two years, he has given the impression of being as indifferent to the need for arms control as during his election campaign three years ago. And although the House of Representatives resolution that there should be a negotiated freeze on nuclear deployment will not tie his hands, the fact that it was passed at all is a sign that those wishing to succeed in November 1984 would be well advised to talk well, and act well, about arms control. So, with luck, even Mr Andropov may be surprised by what happens at the negotiating session beginning next week. A president eager to match Soviet enthusiasm for arms control could be disarming. That at least must be the hope. □



Research without benefit

The British defence establishment has been warned that its research must be more useful to others.

Is there a chance that the next British Government (which will probably take office four weeks from the date of this issue) will find a way of bringing its spending on defence research and development within the compass of its general economic policy? This question is not as partisan as it may seem, for it is no more pointed now than at any of the other elections since the Second World War. For it has been a constant source of complaint since then that defence research and development has taken the lion's share of all spending on research and development, but that British industry has benefited only in minor ways from what has been spent. At the outset, when the momentum of the wartime research establishments was still huge, it was also possible fairly to complain that too great a proportion of what the services spent on research and development was spent within their own research establishments, but over the years the proportion has declined with the enforced restriction of public spending to the point at which most development and nearly half of research in military fields is spent on contracts with industrial companies. Yet even with this huge shift of resources, British industry seems not to be a conspicuous beneficiary of military spending. Why should things be so different elsewhere, in the United States particularly, but also in countries such as France?

For many of the past several decades, successive governments' critics have been asking that there should be a formal inquiry into this important question, but so far in vain. For most of that time, the Ministry of Defence has taken the view that defence research is too vital to national security to be compromised by merely commercial considerations — and that, anyway, it is secret. The nearest approximation to such an inquiry was that which led in 1962 to the Gibbs-Zuckerman report which was so devoid of tangible examples of what was being done in the name of defence research that it has served principally as grist for academic mills. But the climate is obviously changing. The British Government's own Advisory Council on Applied Research and Development is said to have embarked on studies which touch the defence establishment — but will not say what they are about. And the National Economic Development Office, an institutionalized talking shop in which industry, trades unions and the government (the present government reluctantly) participate, had the wit a year ago to commission a study on defence research in relation to electronics from Sir Ieuan Maddock, once chief scientist at the Department of Industry who would no doubt have risen even higher if only more than ten per cent of the British people had known how to pronounce his first name.

As it has emerged, Maddock's report makes one telling point — that the industrial companies most closely involved with defence research are so dependent on their military customers that they do not care, or dare not risk, to spread their wings into other fields unless into various aspects of civil aerospace or the export of military equipment to countries overseas. But the report points to companies at the other end of the spectrum which operate successfully in civil markets for high technology products but which are kept from competing for defence contracts, or even from becoming subcontractors, by the sheer clubbiness of the present system. The report rehearses the many complaints that have been made against the defence research establishments, usually that they attempt too much with their own resources, and comes to the unfashionable conclusion that if civil benefit were the objective, and because the well-established defence contractors are unlikely ever to contribute to civil manufacturing, the defence research establishments should carry out more research than they do at present in the hope that they might then channel some of the benefits towards companies at present outside the defence establishments. This proposal is briskly rejected in a note from the Ministry of Defence published as an appendix.

As things are, the ministry is probably correct. But its devotion to the causes of putting more of its research out to private industry

(and why not universities?) would be more convincing if it had emerged much sooner, when the defence laboratories were allowed to employ far too great a share of the country's skilled manpower.

So what is the best way of exploiting defence research? The simplest answer is probably "none of the above". For Sir Ieuan's report in passing is a sombre reminder of how a shrinking national product has made minuscule what even the defence ministry, the richest customer for research in Britain, now spends. The cost last financial year of the research supporting the British services was merely £300 million, less than the annual expenditure on research and development of many Japanese manufacturers of consumer products. And although the Ministry of Defence spent five times as much on the development of weapons and other gadgets, these days even that would not see a modest development of an under-sized supersonic aircraft such as Concorde to completion. Is it that while the British have wasted nearly four decades on how to make civil use of defence research, the goose that used to lay the golden eggs has become a silver gosling? □

Records in thin air

The Hitler non-diaries should remind historians that their scholarship is not self-sufficient.

THE fiasco of the supposed rediscovery of Adolf Hitler's diaries is primarily an embarrassment for the newspaper groups that have enjoyed the sense of having scooped their rivals only for a few short weeks — *Stern* in West Germany and the *Sunday Times* in Britain. For they must now expect that their readers, given such a spectacular display of gullibility, will ask similarly awkward questions about the rest of what they print. Much the same is true of the academic historians who have been caught up in this tangled tale as "verifiers" of the supposed diaries, most especially Lord Dacre (previously the Oxford historian Hugh Trevor-Roper) who spent four hours with the forged documents in the vault of a Swiss bank and publicly, if only briefly, pronounced them to be "authentic". That the newspapers concerned may have been misled by the ironical dictum that too many good stories are lost by too much checking is understandable if inexcusable. That historians should similarly have been led astray is mystifying.

The proofs of forgery adduced last week by the West German Government are not, as it happens, unfamiliar. Diaries supposedly written by Mussolini were quickly shown to have been written on paper manufactured only after his death. Now there are several technical means by which forensic scientists can tell something of the provenance of writing materials. Mussolini's diaries were discredited by a proof that the chemical and fibre composition of the paper was inconsistent with plausible origins. Since then, X-ray analysis of clay used as a smoothing agent in the manufacture of paper has in some cases proved decisive, principally because the trace constituents of clays vary from one region to another. Similarly, the chromatographic analysis of inks has uncovered other fakes, while even pollen analysis has been used to show that forged documents are false by demonstrating that pollen grains embedded in the ink do not correspond with the season at which the documents are supposed to have been written.

While these techniques are by no means simple, the fact of their existence is familiar and their use is widely practised. Moreover, they are techniques that would have splendidly suited *Stern's* purpose in that it would have been possible to ask whether samples of supposedly pre-war paper were plausibly such without disclosing what was supposed to have been written on it. Putting an historian, however distinguished, in a bank vault with the papers for a few hours is no substitute for objective tests of authenticity, even though it now emerges that many of the entries in the supposed diaries were simply copied from a book published twenty years ago. The next time (and there will be one) that newspapers and historians are required to decide whether documents are forgeries, they should in their own interests give science a chance. □

US national laboratories

Labs to be shielded from governmental whims

Washington

ACCUSING the Department of Energy (DoE) of mismanaging its national laboratories appears to have become the latest blood sport in the United States. A group of private sector consultants told President Reagan last month that the 12 laboratories are an impressive national asset, but one that is hamstrung by too many layers of DoE officialdom and muddled about its role. A review last September by DoE's own Energy Research Advisory Board (ERAB) reached similar conclusions. Now the White House Science Council is about to enter the fray.

The flurry of reports may be bad news for DoE but good news for the laboratories. Both the private sector report (part of the Grace Commission survey of government cost efficiency) and the ERAB report praised the laboratory directors for capable management. They want the directors to be given more control over their own budgets and DoE to relax its detailed supervision and reporting requirements. Both reports recommend cushioning the laboratories from capricious fluctuations in government funding by setting budgets two or three years at a time.

All these recommendations, if implemented, would be a tonic for the laboratory directors. Dr Herman Postma, director of Oak Ridge, told the Grace panel that the number of progress reports the laboratory has to file with its federal masters had doubled over the decade to reach an annual level of 700. Annual fluctuations in congressional funding are an even more serious problem: the Grace investigators reported that in some areas, such as fossil fuels, funding had zigzagged between increases of 67 per cent and reductions of 51 per cent. The result had been a spate of half-completed projects and wasted money.

Less welcome to the directors is the insistence of both reports that DoE should devise a sharper definition of the kind of research the laboratories should do. The Grace report complains that in many cases the laboratories hang on to research projects long after they could have been handed over to industrial sponsors or to universities. It says research on geothermal energy, batteries and turbines can be handled well enough in the private sector. In energy research, the laboratories should look only at areas with big energy potential, such as fossil fuels and nuclear energy — not hydropower, geothermal energy, solar or wind energy.

The White House Science Council report, which is to be published, belatedly, at the end of the month, is certain to

reiterate the call for clearer mission guidelines. One of the main jobs assigned to it by presidential science adviser George Keyworth last June was to reapportion research responsibilities between the federal laboratories, universities and industry. Unlike the Grace and ERAB reports, the White House panel has reviewed the whole gamut of federal laboratories and may therefore recommend some radical change, such as removing the national laboratories from DoE and placing them under a new cabinet department.

But members of the Grace and ERAB groups, who briefed the White House panel, regard that prospect as unlikely. A safer bet is that the report will ask the national laboratories to focus on their traditional strengths in nuclear physics, nuclear medicine, breeder technology and some aspects of the life sciences, while passing on as much research as possible to the private sector and universities.

US nuclear power plants

Emergency planning inadequate

Washington

Two nuclear power plants already operating in the state of New York are to be closed down by the Nuclear Regulatory Commission (NRC) next month because emergency evacuation plans for the local area are considered inadequate. The operators of the plants — both at Indian Point, some 30 miles north of New York City — have until the end of May to improve their plans or persuade the commission to change its mind.

The NRC decision, the first time it has threatened to close working plants because of worries about local evacuation, followed a report by the Federal Emergency Management Agency on a practice drill in March. The agency said two problems — the refusal of nearby Rockland County to join the planning and doubt about the availability of Westchester County bus drivers in an emergency — meant emergency planning was inadequate. About 290,000 people live within 10 miles of the reactors.

Rockland County's refusal to cooperate appears to be financially motivated. Mr John Ahearne, an NRC commissioner, said the county was not served by the plants and received no tax income from them. The cost of participating in emergency planning would exceed the amount provided for that purpose by the state. The participation of Westchester bus drivers in an emergency was doubtful because the county had been

unable to devise an acceptable contract. What impact will the succession of reports have on the national laboratories? Not a great deal, according to several directors. The consensus among those who have surveyed their work is that, despite organizational wrinkles, the laboratories are working efficiently and producing good research. Meanwhile DoE has already begun to look for a sharper definition of the kinds of research that should properly be done by the laboratories and those that properly belong elsewhere.

In February, a report by ERAB split energy technology research into three categories: research in which the federal government should have a primary role, research in which the federal role complemented that of the universities and industry and research in which the federal role ought to be minimal. The only areas selected for the first category turned out to be magnetic fusion, reactors, advanced concepts in geothermal energy, nuclear waste, uranium enrichment and the assessment of uranium resources.

Few of the dozen laboratories will find this division of research roles hard to accept. According to Dr Postma, the problem facing the laboratories is not of defining missions but of achieving some financial stability.

Peter David

unable to devise an acceptable contract.

Spokesmen for the joint operators of the two plants, Consolidated Edison and the New York Power Authority, expressed confidence that any deficiencies could be remedied in time to avert a shutdown. They said that nuclear plants had operated safely at Indian Point for 20 years, and claimed the cost of a shutdown would be enormous.

NRC requirements for evacuation planning were stiffened after the accident at Three Mile Island in 1979 and have not yet been formally met at many plants. At Indian Point, considered by some scientists to pose the most complex evacuation problems in the United States, deadlines for meeting the requirements have been missed on two occasions, and the site has become a test of NRC's seriousness about emergency planning. Mr Ahearne said many people wanted to see whether NRC would stand by its regulations. If it did not, national planning for radiological emergencies would "rapidly deteriorate".

At Indian Point, NRC appears to believe that a satisfactory evacuation plan is feasible given cooperation between the plant operators and the local governments. Mr Ahearne said the operators appear to have devoted too little effort to developing plans with their local government neighbours — perhaps because of a belief that NRC would not close the plant simply because of the absence of a workable plan.

Peter David

Remote sensing

British images up for sale

A £14 MILLION programme to promote development and commercial exploitation of remote sensing in Britain was announced last week by Mr Kenneth Baker, Minister for Information Technology. The immediate aim is to seek "maximum possible benefits" from the European Space Agency (ESA)'s Earth Resources Satellite (ERS1), due for launch in 1987.

A new coordinating body, the National Remote Sensing Programme Board, will run the programme. The board, which includes representatives from government departments, research councils and industrial companies, will advise on the resources needed and will promote applications of remote sensing. As well as overseeing the ERS1 programme in Britain, it will supervise the National Remote Sensing Centre at Farnborough, which processes data obtained through ESA's "Earthnet" scheme. Mr Baker said that this agreement will encourage private companies to exploit added-value services as soon as possible.

Behind the government's enthusiasm there is concern that industry has been slow to make use of remote sensing data. Opinion is united on the aesthetic value of remotely-sensed images, but US experience is not encouraging. When the National Oceanic and Atmospheric Administration took over the Landsat programme on 31 January this year, only one serious customer was found for the data and then only after especially favourable terms (which included government purchase guarantees) were offered (see *Nature* 10 March, p.96). The British Government has, however, declined to follow the French course of helping to set up an independent company to market remote sensing data.

The £14 million set aside for the new scheme is on the low side of the recommendations made to Mr Baker a year ago by a specially convened "task force". The government seems to have recognized the need for development in ground computing techniques to put data into a usable form, but some scientists feel there should be more support for the development of more advanced sensing instruments.

ERS1 will be a partly experimental oceanographic and coastal monitoring satellite along the lines of Seasat. The British contribution to the satellite, which will be launched by Ariane 2 or 3, will be worth £10 million per annum by 1984-85. The main payload will consist of a radar altimeter, which can measure the satellite's height above the waves with a precision greater than 10 cm, and an active microwave instrument combining the functions of a wind and wave scatterometer and a synthetic aperture radar. This instrument will measure wind and wave directions in all weather conditions. An additional

experiment, the Along Track Scanning Radiometer, has been provided by the UK Science and Engineering Research Council out of its own funds.

Apart from their academic interest, data from ERS1 will be used in forecasting weather and sea conditions, including sea-ice, and for advisory services to offshore

industries. Possible future applications for remote sensing include fisheries management and monitoring marine pollution. A land applications satellite to follow ERS1 is planned, although few details have been decided. In commercial terms, land data are likely to be more profitable in the long run, with obvious applications in mineral prospecting and in agriculture. But experience with rapid-turnaround ocean data may make it easier to persuade industry to take notice.

Tim Beardsley

British AIDS

Whose blood can now be safe?

FEARS that the spread of acquired immune deficiency syndrome (AIDS) in Britain may be hastened by the use of blood products imported from the United States have prompted a trades union leader to attack the government for allegedly failing to provide adequate support for the Blood Products Laboratory (BPL) at Elstree, in Hertfordshire.

The worries concern supplies of factor VIII, the clotting agent used to treat haemophilia. There are about 4,000 haemophiliacs in Britain, and 60 per cent of the factor VIII they use is imported from the United States, where (in contrast to Britain) donors are frequently paid to give blood. An official of ASTMS, the union representing technicians at BPL, said last week "It is common knowledge that many blood donors in the United States are on Skid Row and may be carrying AIDS". The disease is thought to have a long latency period.

No causative agent for AIDS has been identified, but there is evidence suggesting that AIDS is transmitted in blood products, (see *Nature* 28 April, p.749). There are 11 haemophiliacs among the thousand or so AIDS victims in the United States — more than would be expected by chance — and, along with male homosexuals, intravenous drug users are also known to be especially at risk. A recent issue of *The Lancet* records the first three reported cases of AIDS from Spain, all of whom are haemophiliacs. There are now 14 confirmed cases of AIDS in Britain, and one recent suspected case is in a male haemophiliac who is neither a homosexual nor a drug user.

A second factory now under construction at Elstree should make Britain self-sufficient in factor VIII by 1985-86, but until then the country is dependent on imports. Mr Clive Jenkins, general secretary of ASTMS, claimed last week that managers at BPL have been told to cut their budgets by 10 per cent. The Department of Health has strongly denied Jenkins' claims, which it calls "entirely untrue". The department says that, in line with other public services, the laboratory's budget has been increased by 5 per cent. But Mr Norman Pettit, assistant to the director at BPL, confirmed last week that 10

per cent cuts had been ordered. It is not clear whether cash limits will delay the plant.

Most factor VIII is now extracted from pooled donations from 4-5,000 people. This makes the potential risk much greater than with ordinary blood transfusions, which typically contain blood from only three or four donors.

Dr Harold Gunson, who advises the British Department of Health on blood products, said that the situation was being kept under review, but that in his opinion there was insufficient evidence to justify reducing US factor VIII imports immediately. There is growing pressure on the National Blood Transfusion Service to exclude homosexuals from its pool of donors, but Dr Gunson says "We just can't do that". It is thought that asking direct questions about donors' sexual proclivities would be unacceptable in Britain — and might not yield the truth.

All such problems would vanish were it possible to produce factor VIII by means of recombinant DNA technology. That could become possible within 18 months (although not necessarily as a commercial venture) through the joint efforts of Dr Edward Tuddenham and his colleagues at the Royal Free Hospital in London, the British blood products company of Speywood Laboratories and Genentech, the Californian biotechnology company. Speywood carries out the first steps in purification of factor VIIIc (the subcomponent of factor VIII that is deficient in haemophilia) which Dr Tuddenham purifies further. Genentech is trying to clone the factor VIIIc gene and to express it in microorganisms. Rivalry comes from Biogen with Kabivitrum and from the Scripps Clinic and Research Foundation with Johnson and Johnson.

Speywood's backers, the British Technology Group and Prutec, have recently insisted that the company should cut back on some of its research expense in universities. Dr Tuddenham's work has apparently not suffered as a result but research parallel to that of Genentech's on factor VIIIc gene cloning under Professor George Brownlee at the Sir William Dunn School of Pathology in Oxford is being delayed for lack of full support from Speywood.

Tim Beardsley

In vitro fertilization**Success breeds more problems***Canberra*

A DEVELOPMENT in the *in vitro* fertilization (IVF) work of Professor Carl Wood's group in Melbourne has resulted in a successful pregnancy following implantation in a woman's uterus of an embryo previously frozen for four months in liquid nitrogen. The success was announced last week by Dr Alan Trounson, who implanted the embryo in an infertile patient 14 weeks ago. Two of the eight cells of the embryo were damaged during thawing but a normal fetus is assured from the multipotential nature of embryonic cells at this early stage.

The team from Monash University's department of obstetrics and gynaecology, at the Queen Victoria Medical Centre, responsible for 42 births from IVF since 1979, have been freezing embryos since late 1980 but previous implantations have not resulted in pregnancy. Attributing their success to a change in freezing technique, Dr Trounson said that more pregnancies from some of the 35 embryos now in cold storage can be expected soon.

Early last year, members of the group were implanting in patients embryos developed from eggs donated by other women, in addition to their usual routine. The Victoria government of the day responded to public concern by instituting a committee to examine the social, ethical and legal issues arising from the work and, if necessary, to recommend legislation. The IVF work has since been under the scrutiny of that committee, chaired by law reform commissioner, Professor Louis Waller.

The committee's first interim report in September 1982 was cautious, considering only the "easy" case — that of husband and wife supplying gametes to be fertilized *in vitro* and implanted in the wife's uterus. The committee found this acceptable and recommended that IVF be limited to these cases, *ipso facto* ruling out the donation of eggs but allowing embryo freezing. Moreover IVF was justified solely as a means of curing infertility. Because of its relatively high cost — about A\$42,500 per couple, largely paid for by health insurance companies — IVF was seen as a last-resort treatment as the report required a couple to "have undertaken all other medical procedures during a period in excess of 12 months which may, in their particular circumstances overcome their infertility". The Waller committee also wanted all embryos to be implanted.

The IVF group responded by ceasing the donor egg programme but continued with freezing embryos. Indeed, storing embryos appeared to be the only way of ensuring that all embryos were in fact implanted, as some could be available for subsequent implantation if the first attempt failed, instead of being wasted. Moreover, the

number of laparoscopies could be reduced, thus sparing the patient considerable inconvenience. According to the protocol developed by the Melbourne group, ovulation is induced by the administration of fertility drugs so that as many eggs as possible are collected in a single laparoscopy. The overall improvement in the number of pregnancies per patient since 1979 is largely due to the increased fertilization rate and the fact that more than one embryo is implanted simultaneously. However, the number of pregnancies per embryo implanted has remained low.

Another document bearing on the IVF work is the first report last August by the Working Party on Ethics in Research of the National Health and Medical Research Council (NH and MRC). This body funds medical research but does not have any power to enforce its rules except by withdrawal of funds. All its projects concerned with human experimentation have to be approved by an institutional ethics committee.

The report's supplementary note on IVF is more liberal than the Waller report and approves of IVF treatment "within an accepted relationship". It also approves of the donor egg programme. However "continuation of embryonic development beyond the stage at which implantation would normally occur" and cloning experiments are forbidden. Only early undifferentiated embryos may be stored, with an upper time limit of 10 years which should not be "beyond the time of conventional need or competence" of the women who supplied the egg. Surrogate motherhood is declared "not yet capable of ethical resolution". The most significant recommendation was the setting up of a National Medical Research Ethics Committee, which has since been established. The role of the committee is to review ethical considerations of medical research and respond to questions referred to it by institutional ethics committees.

Despite increased debate on IVF over the past two years and the Waller committee's contribution, the state government seems to be no closer to formulating legislation. Surprisingly, scientists themselves want legislation, perhaps to protect themselves from litigation but also from a desire for guidance from the community and to avoid incurring public displeasure.

Vimala Sarma

Tim Beardsley adds: In Britain, Dr Robert Edwards, who pioneered the technique of *in vitro* fertilization, greeted the Australian group's announcement as "wonderful news". Several UK ethical committees have now approved embryo freezing techniques (see *Nature* 28 April, p. 739). Dr Edwards said "We must now look at our own position. In accordance with patients' interests we have got to go ahead".

Pauling institute

Lawsuit settled out of court

THE Linus Pauling Institute of Science and Medicine at Menlo Park, California, has agreed to pay \$575,000 to settle a lawsuit brought by Arthur Robinson, co-founder of the institute and for an initial period its president. The suit has been hanging over the centre for the past two years, some time after Robinson had ceased to be director. It has now been settled out of court.

Robinson, in his suit, which had claimed several million dollars, alleged breach of contract by dismissal, slander by way of Pauling's alleged assertion that his work was "too amateurish" to be published and that Pauling had destroyed his data and peptide collections.

Robinson and Linus Pauling differ in their accounts of the effect of vitamin C on the incidence or the course of human cancer. Robinson on the telephone listed among his complaints the allegation that his failure to find a beneficial connection between vitamin C and cancer was suppressed but Pauling, also on the telephone, denied this.

According to Pauling, the lawsuit had been settled out of court only in the light of the trustees' concern about the likely cost and trouble of a protracted hearing. He insisted that the amount of the settlement represented no more than compensation for loss of office and the cost of Robinson's legal fees.

Robinson, who has founded a small private research institute of his own in Oregon ("only three employees and hope", he said), said the money from the settlement will help to get the institute off the ground. He plans to conduct research on nutrition and cancer in mice, and ultimately to address more basic research problems. □

A day for Sakharov

THE Congress of the United States is expected to pass this week a bill to proclaim 21 May (the birthday of Andrei Sakharov) as National Sakharov Day. According to Dr Edvard Lozanskii, a former Soviet dissident and a leading campaigner on behalf of Dr Sakharov, the bill will also authorize President Reagan to appeal to all other nations to proclaim their own Sakharov Days. Vera Rich

ICRF director

AN error in 28 April *Nature* (p. 739) made Dr J. Wyke into director of ICRF Medical Oncology Unit. In fact, Dr J. Malpas will continue to be its director when Dr Wyke becomes responsible for its laboratory work. □

US agricultural research

Worry over research grants

Washington

A SHAKE-UP at the competitive grants office of the US Department of Agriculture (USDA) has raised new concern about the future of the programme which provides \$16 million a year for peer-reviewed research proposals in basic agricultural science.

Under the reorganization now taking place, Dr Holly Schauer, the associate chief of the programme who is widely viewed as a strong advocate of its scientific and budgetary integrity, has been forced out, and several largely administrative offices are being combined with the competitive grants office under a new director, to be named shortly. The addition of new administrative duties to the director's job raises the possibility that the job may go to a USDA career bureaucrat, rather than an academic researcher, as has been customary. The directors have until now served for one year. There is growing speculation that the job may go to Dr Clarence Grogan, a USDA career official who has served as acting director of the office since last fall.

These changes have supporters of the programme sitting on the edges of their seats. Since the establishment of the programme in 1978, it has come under constant attack. Opposition has been particularly intense from the land-grant colleges, which receive automatic funding under a state-by-state allocation formula dating back to 1887. Although the competitive grants programme has never had more than a small fraction of the funds that go to the land-grant colleges (about \$150 million), it has been seen as a constant threat.

Unlike the formula funds, the competitive grants are open to all comers — notably researchers at non-land-grant institutions such as Harvard and Stanford universities — and are scrutinized by peer-review panels.

According to Dr Ed Kendrick of USDA, the director of the new office will deal mainly with administrative matters. Advertisements for the position specify a bachelor's degree at least; Kendrick says the major qualifications will be scientific background, scientific understanding, credibility in the scientific community and management skills.

Answering criticisms that this would lower the credibility of the programme and weaken its links with the scientific community, Kendrick said a chief scientist will be appointed, and he stressed that the scientific oversight provided by this person and by the peer-review panels would not be subsumed by the administrative director. He said the reorganization was in fact designed to raise the visibility of the programme by putting it on an equal footing with the Agricultural Research Service

(USDA's intramural research organization) and the Cooperative State Research Service (which administers the formula funds).

Schauer, who is leaving her position on 13 May, argued, however, that the "visibility and integrity" of the programme came from having a respected academic researcher as its director. The changes will mean that "we won't have someone whose main responsibility is the competitive grants programme".

And Lawrence Bogorad, a plant scientist at Harvard University, voiced concern that without Schauer, and without an outsider in the director's chair, the programme may be more vulnerable to traditional bureaucratic manoeuvring. "If you get a real inside person, he can't fight for the money", Bogorad said, whereas an outsider, who has no stake in career advancement within USDA, has nothing to lose in fighting and stepping on toes if necessary.

Stephen Budiansky

Seveso trial begins

THE Seveso trial will have begun in earnest this week (11 May) at Monza, near Milan. Although the trial opened as originally planned on 18 April, it was promptly adjourned after a number of residents from the area of Seveso sought to join to the criminal trial of three employees of the operating company civil suits for the recovery of damages. One possibility is that the court at Monza may decide that these should be dealt with separately, another is that the claims may be settled privately.

The Italian company whose employees are being prosecuted is Givaudan, a wholly owned subsidiary of Hoffmann La Roche. The company's Meda plant near Seveso had been built for the production of trichlorophenol, itself an intermediate in the manufacture of the antibacterial agent hexachlorophene. The accident in which an estimated 150 grammes of dioxin was released occurred on 10 July 1976, when 18 square kilometres downwind of the plant were contaminated and close on 750 people were evacuated from the most seriously contaminated area.

For practical purposes, Roche has already admitted liability for many of the consequences of the accident, and had by the end of last year set aside £46 million in payments to inconvenienced local authorities and for contamination and had paid £11 million in compensation to individuals, including farmers and people whose homes or land were contaminated. Whether the managers and designers of the plant were criminally negligent will be decided by the Monza court, but not quickly. □

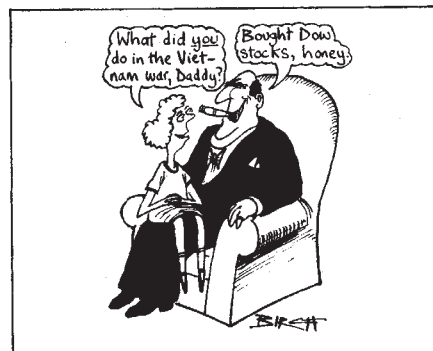
Dioxin

When was the danger known?

Washington

IN the United States, where more than 20,000 Vietnam war veterans are suing the chemical industry for illnesses and genetic damage they blame on exposure to dioxin within the defoliant Agent Orange, a federal court has been hearing a series of spectacular charges and counter-charges about the point at which the industry and the federal government became aware of the health dangers of dioxin.

Lawyers for the veterans have suggested that the industry, led by the Dow Chemical Company, knew as early as 1965 about the toxic impurities within 2,4,5-T — a principal component of Agent Orange — but suppressed the knowledge to prevent "excessive" government regulation. Dow, meanwhile, has alleged that the Depart-



ment of Defense was aware of evidence linking dioxin with birth defects several years before using Agent Orange in Vietnam.

Court papers filed for the veterans include a memorandum written in March 1965 by a chemist for the Hercules Powder Company shortly after attending a private meeting convened by Dow at which company scientists reported that toxic impurities in 2,4,5-T had caused severe liver damage in rabbits. In a memorandum several weeks later, a fellow Hercules scientist reports having been telephoned by a Dow executive and warned to keep the findings away from the federal government.

Now Dow has alleged that both the company and the Defense Department were in possession by 1969 of a National Cancer Institute study linking dioxin with birth defects in mice. Spraying with Agent Orange in Vietnam did not end until 1971. If Dow can satisfy the court that the federal government knew of the dangers, and that Dow merely produced the herbicide according to government specifications, it cannot be held liable for damages sought by the veterans.

Until now, Dow's defence has been based on its belief that exposure to the herbicide could not have been responsible for the injuries the veterans ascribe to it.

Peter David

Soviet science centre

Leningrad's energetic future

A NEW science centre is to be established in Leningrad, Academician Gurii Marchuk, chairman of the Soviet State Committee for Science and Technology of the USSR, announced recently. During the past 25 years, a number of science centres have been established in the Soviet Union, the best known being the Siberian Branch of the Academy of Sciences of the USSR, based in Novosibirsk, which Dr Marchuk himself headed for several years. During the last two quinquennia, the establishment of such centres has figured prominently in the state plans, and has generally involved little more than the formal linking of existing research facilities in the area with local industry, in order to build up the research and development base of the latter.

Leningrad, however, has a long scientific history — indeed (as St Petersburg and then Petrograd) it was until 1934 the home of the Academy of Sciences, and it still houses many important scientific institutions.

The new centre will bring together all the institutes in and around Leningrad and will be aimed at developing not the whole science base but a particular sector — energy. In the Leningrad region, there exists a concentration of energy-related design bureaux, factories and institutes, fifteen of which, Dr Marchuk said, have already been united into “virtually a complete complex”. The designing of radically new energy equipment will therefore be the most important subject tackled by the new centre.

Considerable work has already been undertaken in this field. Long-term plans for the Leningrad region (up to the year 2005) include the development and construction of a thermonuclear (fusion) power station, which, it is hoped, may be commissioned before the end of the century.

Another innovation is a “cryoturbogenerator”, employing the principles of superconductivity, with a promised efficiency of 99.5 per cent. An experimental model is already reported to be in operation, and a full scale model (300 MW) is promised by the end of 1995. Of less theoretical significance but still a major innovation is the production, by the Izkovskii works (in conjunction with the Atom-mash plant in Volgograd), of the first nuclear power station for urban heating, which will be sited in Gorkii on the Volga — a city of 1 million inhabitants. The construction of 12 such urban heating plants promised for the cities of the western parts of the Soviet Union before the end of the century does not, apparently, cause any alarm to Soviet environmentalists.

Such a nuclear district heating system will eventually serve not only the four million inhabitants of Leningrad but also several industrial enterprises.

In addition to the manufacturing plants which produce generating equipment — *Elektrosila* and the Leningrad Metal Works among others — the area also houses *Gidropromekt* which carries out major surveys before construction of hydro-electric power stations (its research department recently developed an encapsulated television camera for lowering into test boreholes), and the Vyborg shipyard which has now started production of drilling rigs for use in the Baltic (previously these were constructed only at Astrakhan for use in the Caspian Sea). The All-Union Institute of Halogens and the Leningrad Pigment Association. These latter two bodies, although not apparently energy related, have it appears an important role in the energy sector — producing anti-corrosion and waterproofing materials for the underground structures of power stations.

How significant the new science centre will prove to be in the general context of Soviet science is difficult to judge. Formally an organ of the Academy of Sciences, it was not considered of sufficient importance to merit a mention in the *Pravda* reports of the annual general meeting of the academy last March. During the past 50 years, Leningrad scientists have felt they are no longer at the true centre of things, that Moscow is the true focus of Soviet science. The new centre is promised a special science village in the Shuvalo suburb of Leningrad. How far the new, utilitarian, centre will satisfy the aspirations of Leningrad scientists for a more central place in Soviet science remains in some doubt.

Vera Rich

US space programme

NASA will wait and see

Washington

THE National Aeronautics and Space Administration (NASA) has decided to postpone launching the second of its sophisticated new communications satellites until it discovers why the first, launched last month by the space shuttle Challenger, failed to arrive in its proper geosynchronous orbit. Postponement of the launch, originally scheduled for August, will impair but not stop the first flight of the European Space Agency's Spacelab a month later.

Plans for the Spacelab mission were based originally on the assumption that both new TDRS (Tracking Data and Relay Satellites) would be fully operational in time to handle the large volume of data generated by the 24 European and 13 United States experiments Spacelab will perform. NASA and the European Space Agency have nevertheless decided that Spacelab's 30 September mission will be “scientifically viable” as long as one TDRS

Perpetual motion?

“STILL it moves!” Galileo is supposed to have muttered 350 years ago, after his imprisonment by the Inquisition for his belief in the moving Earth and other impious doctrines. “Still it moves!” one could say on Monday, when Pope John Paul II delivered a most cautious and uncommitted speech on the subject to a meeting of scientists, including 33 Nobelists, in Rome.

In a piece of classic prevarication — no doubt enforced by ultra-conservative elements in the Church — the Pope announced the establishment of a committee,



Professor Zichichi and Pope John Paul II

an “interdisciplinary research team for the careful study of the whole question”. There are good grounds, said the Pope “for hoping that it will make an important contribution to the examination of the whole question”. Still it moves, still it moves!

The meeting, addressed by Pope John Paul II, was convened by Professor Antonio Zichichi, president of the Italian institute for nuclear physics, in honour of the Galileo anniversary, and will continue for a day and a half to review progress in science since 1633. Clearly Professor Zichichi will have to continue longer with his efforts to persuade the Church finally to rehabilitate the “father of science”. Robert Walgate

is operating properly.

It is still possible that the second TDRS will be launched in August as planned, provided that NASA is confident that there will not be a repeat performance of last month's mishap. Unless the second satellite is launched, however, Spacelab's September mission will depend on NASA's success in manoeuvring the first TDRS into its proper orbit. A series of engine firings aboard the TDRS was due to begin last week in order to shift the satellite out of its present elliptical orbit.

A successful re-alignment of the TDRS will enable the shuttle orbiter Columbia to carry Spacelab into position in September but the absence of a second operational TDRS will reduce the scope of the experiments Spacelab will perform. A spokesman for the European Space Agency said that no experiments would have to be cancelled, but some would produce about 25 per cent less data than planned.

Peter David

Soviet biotechnology

Science confused by politics

A CONFIDENT self-appraisal of Soviet biotechnology was given in London last week by a delegation from the Soviet Academy of Science to a biotechnology conference and exhibition. The delegation, led by the academy's vice-president Academician Yuri Ovchinnikov, also rehearsed for a tiny press conference in London a fierce attack on United States policies on defence and other matters.

According to Ovchinnikov, the Soviet academy began planning a programme of research in biotechnology as long ago as 1972. Since 1974, there has been a formal plan for collaboration between academy institutes which was concentrated, to begin with, on traditional biotechnology. But now, Ovchinnikov said, both insulin and human growth hormone are being manufactured from genetically manipulated bacteria.

The Soviet strength, according to Ovchinnikov, is in microbiology, where the Soviet Union is "ahead" of even the United States. Last week, however, he acknowledged the expertise of West Germany in the manipulation of plant genes, now one of the chief objectives of the academy's programme in biotechnology designed as a contribution to increased food production in the Soviet Union. The use of materials related to pheromones as specific pesticides, of cellular manipulation for the production of new plant hybrids and of embryonic transplantation for multiplying the productivity of superior breeding females are among the immediate goals.

The rights and wrongs of the CoCom strategic embargo arise even in this connection, and the delegation complained that it was "foolish" that the sale by the West of many of the materials required in genetic manipulation is embargoed when the Soviet Union is self-sufficient in radioactive chemicals, restriction enzymes and the like.

By way of answer to a request from a Soviet journalist for a denial of the claim that the Soviet Union lags behind the United States in microelectronics, Ovchinnikov said that the Soviet Union would not have been able to launch sophisticated spacecraft if the *canard* were true, that both software development and semiconductor research were "in good shape" but that there were difficulties of "organization" in the industrial application of computers.

Academician G.K. Skryabin, chief scientific secretary of the Soviet Academy and an impassioned speaker on general themes, interjected at this point the sensible remark that "no one country can have equally good basic science in all fields". Unfortunately, perhaps because the Soviet and Western concepts of what press conferences are for do not coincide,

the proceedings at this point lapsed into a somewhat pantomimic and certainly unstructured discussion of world affairs.

On its planned discussion with the Royal Society of the Anglo-Soviet exchange agreement, for example, the delegation hinted darkly at the "political" influences that had two years ago reduced the quota (measured in man-months per year) of facilitated exchange and a little reluctantly said that it would like to add to the existing programme some means by which there could be a directed exchange of people between government laboratories. □

Taking a broad view

THE Soviet delegation's discussion of nuclear and other weapons, advertised in advance by the distribution of copies of an appeal for nuclear disarmament by a large group of prominent Soviet scientists (published on 9 April by the Soviet news agency TASS), was introduced by the observation that although it might be held that scientists should not engage in politics, "you cannot ignore what is going on in the world".

The 9 April declaration is a comment on President Ronald Reagan's "star wars" speech of 23 March, advocating the development of new kinds of antiballistic missile defences. To the suggestion that Mr Reagan's musings hardly justified such a strong stand by the Soviet Academy, given the scepticism of the US Congress and of critics outside the government, the Soviet delegation replied that nobody could expect the Soviet Union to count on good sense prevailing, that the Pentagon was already working on novel anti-missile weapons systems and that the Soviet Union would have no chance but to follow suit. According to Skryabin, "we are scientists but are being forced into being politicians" by the threat of nuclear war.

Many of the signatories of the Moscow declaration of 9 April are academicians, for which reason it was natural to ask why Andrei Sakharov had not been asked to sign. Academician Skryabin's impassioned reply was that Soviet scientists did not wish to see their signatures alongside that of the only one among them who had publicly appealed to the US President not to supply grain to the Soviet Union, "thus depriving our children of bread". Mr Skryabin went on to say that Soviet scientists were resentful that Sakharov had all along been paid a salary of 1,200 rubles a month, and that he had "used for his own purposes" of protest the excuse that the Soviet Government had refused permission for his stepdaughter to emigrate, forgetting that her parents had also refused permission. In the circumstances, Skryabin "underlined", Soviet scientists were entirely justified in ostracizing Sakharov.

Soviet human rights

One-way trip for Sakharov?

Jerusalem

THE Soviet authorities may be rethinking their attitude on Dr Andrei Sakharov. According to Professor Peter Weinzierl, head of the Institute of Experimental Physics of Vienna University. Citing "unidentified diplomatic sources", Professor Weinzierl said last month that there is "a great possibility" that the Soviet authorities would permit Sakharov to travel to Vienna to take up a visiting professorship at the university. The Austrian Ministry of Science and Research, which has forwarded the university's invitation to the Austrian Embassy in Moscow, says that there should be no "over-optimism and speculation at this stage".

Similar caution was expressed at the Weizmann Institute at Rehovot, Israel, where on 28 April Sakharov was awarded *in absentia* an honorary doctorate in recognition of his achievements in physics and his services to human rights. A small exhibition at the Weizmann Institute mounted to celebrate the award by Dr Edward Trifonov (a personal acquaintance of Sakharov) and Dr Harry Lipkin (a physicist whose scientific interests closely parallel those of Sakharov, and who represented him at the degree ceremony), drew attention to a statement by the Norwegian human rights activist Viktor Sparre that in 1975 Sakharov was not permitted to travel to Stockholm to collect his Nobel Peace Prize, on the grounds that he was in possession of classified information.

Sakharov himself, moreover, feared that if he left the Soviet Union he might not be permitted to return (as happened to his fellow scientists Valerii Chalidze and Zhores Medvedev), and when applying for a passport in 1975 he had asked for guarantees that he would be permitted to re-enter the Soviet Union. Similar fears prevented him from applying to accompany his wife to Italy in 1977 when she travelled there for an operation on her eyes, since at that time he felt it his duty to remain in Moscow as principal spokesman for the human rights movement.

Since January 1980, however, when he was sent into "internal exile" in Gor'kii, he has lived in virtual isolation, and although in his first months there he managed to do some theoretical work, the lack of access to a scientific library and the repeated loss of his notebooks to security searches have, according to recent reports, made him virtually abandon the attempts.

In these conditions he might well see the invitation to Vienna as his last chance to resume his scientific and humanitarian activities providing, of course, that the Soviet authorities are really willing to let him leave.

Vera Rich

US-Soviet scientific exchange

Human rights still a block

Washington

SCIENTIFIC relations between the United States and the Soviet Union, already at their lowest ebb since the end of the Cold War of the 1950s, may suffer a new setback next month when two of the eight surviving scientific exchange agreements between the two countries become due for renewal.

The State Department is in the throes of a review of the two agreements — one on atomic energy and the other on transport research — and has asked the Departments of Energy and Transportation for an assessment of the benefits of renewal. In both cases, it is believed, the government scientists involved have recommended renewal. A final decision, however, will be based on diplomatic considerations.

Exchanges between the two superpowers are currently running at about a fifth of the level they had reached before the Soviet incursion into Afghanistan in 1979 and the Carter Administration's imposition of

laboratory near Serpukhov, of a new 3-TeV fixed target proton synchrotron which is planned to have a 13-mile circumference. The laboratory already has a major accelerator of 76 GeV. Dr Bernard Hildebrand, chief of physics research at the US Department of Energy, said the new machine would be one of the leading facilities in the world.

One particularly active area of collaboration is in earthquake research. Perhaps surprisingly, the Soviet Government has allowed a team of geologists from Columbia University's Lamont-Doherty Geological Observatory to maintain a network of seismographs around the Nurek Dam as part of an investigation into reservoir-induced earthquakes. Dr David Simpson, a member of the team, said the research was of great importance to the United States and the

Soviet Union had never raised security objections to the presence of the seismographs, which have a range of 10 to 15 kilometres.

The fate of the two agreements due for renewal next month will be closely watched by the National Academy of Sciences, whose members are deeply split about the future of relationships with Soviet scientists. In a recent speech to the academy its president, Dr Frank Press, said the moratorium on relations was "unsatisfactory". It had not changed the attitude of the Soviet Government on the human rights issue and it had deprived scientists of both countries of useful contacts. In a number of fields, he added, Soviet science was at the forefront of knowledge, and many US academy members wanted to maintain connections with the Soviet Academy. "But our members feel strongly about the human rights issue and insist that any agreement must be modulated by Soviet progress in the human rights area."

Peter David

BRL and Gibco may merge

Washington

BETHESDA Research Laboratories (BRL), a major supplier of products used in biotechnology research, is discussing a possible merger with Dexter Corporation's Gibco Division. Earlier this year (see *Nature* 24 February, p.646), BRL let it be known that it was considering a public stock offering. Company officials said then that they felt BRL had recovered sufficiently from its troubles of a year ago (when, faced with a debt of \$9 million, the company was forced to lay off half of its

450 employees) to go public. Dexter and BRL said they would consider a stock offering if the merger goes through.

Dexter, a Fortune 500 company based in Connecticut, produces tissue-culture products and small-volume parenterals and diagnostics. "We have an opportunity to create one of the industry's largest biotechnology companies", said Frederick Adler, chairman of BRL. According to BRL, the new company's sales would approach \$100 million per year.

Stephen Budiansky



sanctions. Last year, the exchanges were pared down again as part of US retaliation for the Soviet role in the imposition of martial law in Poland. The State Department decided at that time that three out of eleven agreements — would be allowed to lapse.

The atomic energy agreement and the transport agreement pose a more difficult problem for the department. In transport, collaboration between the two countries focuses on avionics and air safety. The scientific benefits to the United States are regarded as minimal but Soviet and Eastern Bloc cooperation are essential for the United States to get its way in international air safety organizations.

In atomic energy, government scientists believe that the agreement with the Soviet Union confers genuine scientific benefits, particularly in research into the fundamental properties of matter and fusion science.

A team of American physicists has just returned from a visit to the Soviet Institute of High Energy Physics where they saw the first stages of construction work, at a

Nature index of biotechnology stocks

12-Month high	12-Month low	Company	Close previous month	Close 28 April	Change
76¼*	16½	A.B. Fortia (Sweden)	58	76¼	+18½
23¼	18¼*	Biogen (USA)	21¼	18¼	-3½
8	2	Bio-Logicals (Canada)	5½	5¾	-¾
13	3¾	Bio-Response (USA)	11	10½	-½
16	7¼	Cetus (USA)	15¾	16	+½
14½	6½	Collaborative Research (USA)	13½	10 5/8	-2 5/8
34 *	5¾	Damon (USA)	33½	34 5/8	+1 1/8
37½*	8¼	Enzo-Biochem (USA)	34	37¼	+3¼
28	6 5/8	Flow General (USA)	12 5/8	12 1/8	-½
69	26	Genentech (USA)	54¾	41	-13¾
12½	2¼	Genetic Systems (USA)	10½	10¾	+¼
18	7	Genex (USA)	17½	15 5/8	-1 7/8
27¾	9 5/8	Hybritech (USA)	22	27¼	+5¼
18½	5	Molecular Genetics (USA)	17½	16¾	-¾
27	8	Monoclonal Antibodies	20½	17½	-3
57¼*	34¾	Novo Industri A/S (Den)	51	57¼	+6¼

Closing prices are for the last Friday of the month. Yearly highs and lows are based on Friday trading prices. For over-the-counter stocks, the bid price is quoted; for stocks on the American and New York exchanges, the transaction price. Data from E.F. Hutton, Inc. The *Nature* Biotechnology Stock Index was expanded this month to incorporate Biogen NV, the company headed by Walter Gilbert that went public on 22 March. Weighting factors were recalculated as of 25 March. On 29 April, the index stood at 213, compared with 200 a month earlier.

*Indicates a new high or low set during the past month.

Polish dismissals

SIR — I think your readers will be interested in some recent news about science in Poland. I refer in particular to the situation at the Instytut Badan Jadrowych (Institute of Nuclear Research) in Warsaw, which is a large research institution (staff numbers about 3,000) of the Polish Atomic Agency. The latter authority has recently started a programme of major changes in the structure of the institute, which in a few months should lead to its dissolution and to the creation in its place of the following three research centres: (1) Instytut Problemow Jadrowych (Institute of Nuclear Problems); (2) Instytut Energii Atomowej (Institute of Atomic Energy); (3) Instytut Chemii i Techniki Jadrowej (Institute of Nuclear Chemistry and Technology).

Such structural changes are preceded, these days, by a wave of dismissals of scientists and technicians, the official reason being redefinition of the scientific aims in view of the splitting of the institute. So far, about 15 members of the staff, among them senior and internationally well-known physicists, have received official notes of dismissal, and reliable leaks suggest that this number is going to increase several fold in the near future. All scientists dismissed so far have been actively involved in the past in Solidarity or in related self-management activities, which might induce one to think that the planned structural changes are being exploited by the Polish regime as a means of putting political pressure on them. They are now planning legal action against the dismissals, and they hope their opposition may be strengthened if the Western scientific community shows signs of interest in their case. Within the Polish scientific community it is felt that this episode may be the beginning of a more generalized wave of dismissals of physicists in universities and other scientific institutions, designed to deter political dissent among scientists.

F. PERSICO

*Istituto di Fisica
dell'Università degli Studi,
Palermo, Italy*

DNA polymorphisms

SIR — The term "polymorphism" is used routinely as a synonym for "variant" in the recent literature describing DNA restriction fragments obtained from genomic DNA, regardless of their frequency. For example, a recent paper in *Nature* describes a "DNA polymorphism" which was observed in members of a single family with an inherited disorder¹. This new practice ignores the honourable history of the term polymorphism in the field of population genetics, where it is reserved for variation which is present at substantial frequency in a specific population. Harris² suggests that a useful definition of a polymorphic locus is one for which there are at least two common alleles with frequencies above 0.01. An allele which is pre-

sent at less than polymorphic frequencies in a population is simply designated a variant. I believe it would be useful to retain this distinction between polymorphisms and rare variants. Restriction fragment variations which are present at substantial frequency in a population can accurately be referred to as DNA polymorphisms, or even as RFLPs (restriction fragment length polymorphisms³) if one can pronounce it. However, when rare variants are described, as in the cases of restriction fragment variants associated with rare genetic diseases, the term polymorphism should be avoided.

MIRIAM H. MEISLER

*University of Michigan,
Ann Arbor, Michigan, USA*

Newton telescope

SIR — It is one thing for embryo sociologists at the University of Sussex to practise their craft on imaginary failures in the management of science in the recent past, but quite another for *Nature* to propagate (*Nature* 17 February, p.56) such training exercises. A professional research project would have found, for example, that the two strongest contemporary arguments advanced for siting the 2.5m telescope at Herstmonceux were (1) modern astrophysics involves large telescopes and the approach to astronomy by the new generation was being conditioned by the small reflectors, mainly of amateur origin, then available to them and (2) the British astronomical establishment itself needed training in the operation of a modern observatory, complete with a large reflector. Few would argue that the 2.5m and the Royal Observatory have not served this purpose and ensured the success of later and larger instruments on distant sites.

OLIN J. EGGEN

*Association of Universities for
Research in Astronomy,
La Serena, Chile*

Lead is a hazard

SIR — In the leading article "Looking for a lead" (*Nature* 21 April, p.641) you state that "for most people in Britain, lead in diet . . . is the chief source of lead contamination, with urban and industrial dust for many people a close second . . ." and ". . . if dust is the chief source of lead contamination in infancy, should not the battle against lead in dust deserve equal importance with that against lead in petrol?"

Petrol combustion is the largest source of atmospheric lead pollution¹ and it is reasonable to assume that the bulk (more than 90 per cent) of lead in urban dust is petrol-derived^{2,3}. In addition, it has been shown that close to roads the lead in vegetation is isotopically similar to that in petrol⁴ and that atmospheric deposition accounts for 90–99 per cent of the lead content of grass in remote rural areas⁵; no doubt such levels also apply to crops.

Thus the banning of lead additives in petrol should reduce both dietary lead intake and urban dust lead levels to the benefit of all sectors of the population.

DAVID COLEMAN

London SW15, UK

1. Nriagu, J.O. *Nature* 279, 409–411 (1979).
2. Day, J.P., Hart, M. and Robinson, M.S. *Nature* 253, 343–345 (1975).
3. Smith, W.H. J. *Air Pollut. Control Assoc.* 26, 753–766 (1976).
4. Chow, T.J. *Nature* 225, 295–296 (1970).
5. Tjell, J.C., Hovmand, M.F. & Mosback, H. *Nature* 280, 425–426 (1979).

● The Royal Commission report says that "while the results of our calculations have considerable margins of error, we believe that most adults derive up to 20 per cent of blood lead from petrol, but the figure may be considerably higher for particular groups of people". The article complained of also drew attention to the special case of children. — Editor, *Nature*.

No-lead engines

SIR — We are pleased to see the comment in your notice of the report of the Royal Commission on Environmental Pollution concerning the document submitted by this organization which discusses engines not needing high octane fuels. Although you describe us as "unidentified" (*Nature* 21 April, p.643) we are in fact a group of independent consultants specializing in internal combustion engines and fuels.

We believe that the points made in our document are of considerable significance because by the application of advanced technology to the further development of certain engine forms some of the costs in higher fuel consumption that will arise from the abolition of lead from petrol can be retrieved. Until now the ease with which octane numbers could be increased by adding lead to petrol has been a disincentive to such research.

It would be unfair to expect the UK motor industry to shoulder the whole burden of developing new engine forms. If government decrees that lead should go then it seems reasonable that government should support an urgent programme to ensure that the total cost of the results of that decree be minimized.

Outside the motor industry there are several centres of excellence in this country where government funded fundamental work on new engine forms could effectively be carried on. These include certain universities, the Department of Industry research establishments, the Atomic Energy Research Establishment at Harwell and contract research companies such as Associated Octel Ltd and Ricardo Consulting Engineers PLC.

RICHARD W. WHEELER
MARTIN H. HOWARTH

*Atlantic Research Associates,
Tunbridge Wells, Kent*

1. Karanthas, S.K., Norum, R.A., V.I. Zannis & Breslow, J.L. *Nature* 301, 718–719 (1983).
2. Harris, H. *Principles of Human Genetics*, 3rd Edition, p.311 (Elsevier, Amsterdam 1980).
3. Botstein, D., White, R.L., Skolnick M. & Davis, R.W. *Am. J. hum. Genet.* 32, 314–331 (1980).

Can order spring from chaos?

Italian science, to which the following twenty pages are devoted, is a curious amalgam of ingenuity and muddle and as such a reflection of the political system under which Italy labours.

MAXWELL'S demons tried to create order in a gas by keeping fast molecules on one side of a barrier and slow molecules on the other. So in Italy do a few scientists struggle to create order — a little science — in something near political, infrastructural and financial chaos. The only difference is that thermodynamics condemns Maxwell's demons to failure, while scientists in Italy sometimes succeed.

Many things contribute to the disorder, so many that Italy begins to look like a microcosm of world problems, rather than a single nation. Italy has many divisions. It is a comparatively new country living among conspicuous ruins of the past, which contribute to a feeling of decadence. Local loyalties are strong, stemming from a belief in the individual and a disrespect for structures: for example, Italy has ten times the number of businesses per 100,000 employed than West Germany, and a large fraction of the economy is "black", unregistered and unknown to the taxman. (The average West German company has 80 employees, the average Italian eight.) There is loyalty to the family, to village, to region and then to north and south.

Garibaldi unified an Austrian-influenced north and a Spanish south in 1861, but many northerners still consider him to have been a disaster, saddling an organized country — the northern productive plain — with an undeveloped liability in the south. Seen from the north, the peculiar, literally intriguing character of Italian politics is a southern phenomenon; from Milan, the industrial capital, Rome, the political capital, belongs to the south. Reciprocally, the south looks suspiciously on the north as might a colony on its colonists. From that point of view, 1871 meant that the south had lost a civil war so that the southern bourgeoisie — and the universities — with their Bourbon, Spanish culture, have resisted northern influence as a kind of self-defence. To begin with, southern universities were strong in the humanities but relatively weak in science and when the industrial revolution arrived (late) in the north, the differences were exaggerated.

Now there are a few important poles of science in the south, and there are a few outstanding southern scientists, but the history of the past century still colours the pattern of Italian science. Thus the south absorbs scientific cash without showing corresponding productivity: and until the

right structures and laboratories are in place (and a few institutions are pioneering in that regard) to exploit properly the talent that must be there, it will continue to be a drag on Italy's scientific economy. Moreover, making a bad situation worse, the central national institutions for distributing what small research funds there are (about one per cent of gross national product, half that of most industrialized countries) seem unable to do so effectively — with selectivity — because of the importance of individual power as opposed to objective assessment. Thus while strong good scientists do well, so too do strong second-raters, while politically weak good scientists are lost.

Another related problem has been an old-fashioned attitude to professors and their research: that research is an individual thing, to be done by a man and his personal institute, like a Renaissance artist and his school: a hopelessly rigid system for most modern science. Fascism, and the 'idealistic' philosophies that preceded it, also appear to have done their share of damage. (A certain Mr Gentile, minister of education in the early part of this century, is said to have told his son that experiment was unnecessary, that a few days thought would solve any problem he might have. Gentile aided the collapse of scientific institutions that a more far-sighted Garibaldi had helped to establish.)

The history of even the recent past, though, is not that of muddle unrelieved. Starting with the development in the 1950s of stereochemically-directed catalysts (by Giulio Natta), the Italian petrochemical industry was for more than a decade the spearhead of an "economic miracle" whose successes were nevertheless played out long before the increase in the price of oil in 1973. The enduring benefits that might have been expected from those heady days do not, unfortunately, include arrangements for the organization of research and development freed from the petty encumbrances of local politics and the national bureaucracy. Perhaps the most substantial lasting benefit is that the post-war understanding between Italy and Japan which has the effect of limiting Japanese imports still persists. Japanese car imports are only 2,200 a year, colour televisions to a value of less than \$50,000 a year, the European Economic Community's rules notwithstanding. In the short run, this protection is probably an

economic advantage even if, in the long run, it may blunt the spur to a resurgence of the economic miracle.

So how long can Italy go on muddling through into the high-tech 1980s with its scientific institutions in the mess they are in? The problem was amplified only last week when President Alessandro Pertini admitted the collapse of the 43rd Italian government since the end of the Second World War, and called for early elections. The collapse was precipitated by the Socialist Party, a minor coalition member, which argued that the present government had achieved all it could, that its usefulness was past. Not so. Major and crucial reforms of two Italian scientific institutions, the national research council (CNR) and the national institute for nuclear physics (INFN), were under way, and had just reached parliament. Now it seems they may be nipped in the bud.

The reforms are fundamental for many reasons, some of which are discussed in the following pages, but most of all because the careers of scientists in the two organizations hardly deserve the name. Scientists and technicians enter at the bottom and proceed by seniority to the top, with no promotion by merit and no possibility to inject talent from other institutions (such as the universities) at any level other than the bottom. Salaries are also poor. In these conditions, people are demoralized: they and their science suffer.

That is not to say that the planned reforms would have answered all the problems — but they would have gone a long way to help. One of the reformers, who until last week was serving on the *fifteenth* governmental committee to consider the reconstitution of CNR, says he considers this election "disastrous" because in his eight years of work on the committee and its predecessors "we had never been so close" to success. Very strong opposition from the universities (who wished to retain power over CNR) and the unions (who opposed meritocracy) have now largely evaporated. The remaining obstacle was the tangled process of Italian government. This now seems to have dashed hopes of early success. Italian science can ill afford the loss, if indeed, as research ministers are lately eager to argue, Italy is to research its way out of the recession. □

This supplement has been compiled by Robert Walgate, *Nature's* Chief European Correspondent.

A complicated country

Finding your way around

ITALIAN science rises willy-nilly amid Roman ruins, renaissance glories, peeling honey plasterwork, and what an economist might call structural chaos. The following pages attempt to tread some kind of sensible path around the complicated scenery.

We begin with money (this page) because money and its distribution are what most people worry about most of the time. But a lira on paper is not a lira in the pocket: the article explains why.

Then we turn to the national research council (CNR), which according to its statutes advises the government on science policy — but also funds the universities and does an increasing amount of research itself. But CNR seems confused about its role, and is in a terrible mess with the careers of its scientists since it took a wrong turning in 1975 and became a *parastato* (civil service and bureaucratic) institution.

This same straitjacket confined INFN (high energy physics, page 119) and CNEN (now ENEA, nuclear and renewable energy, page 121) in the same year.

Before 1975, the most-troubled scientific institutions in Italy were certainly the universities, for which the key date is 1968–69 — the year of the student revolution, which went further in Italy than in any other country in Europe. But other troubles struck in 1972, with the blocking of all promotions, a little legislative slip-up that lasted a decade. We cover the universities on page 116.

For readers who might like to start with a personal view, on page 125 Professor Francesco Blasi of Naples sums up some of the main frustrations facing an Italian scientist. One of Blasi's main complaints is the complete lack of a system for training scientists — no PhD system (Italy's first such system is starting this year), and no fellowship scheme, for fear that fixed contract workers will always use legal means to lever themselves into permanent jobs — it being generally accepted in industry that once a person is in a job for a year, he or she has a right to it for life.

The first pages of the supplement thus cover most of the main institutions affecting science, and their problems: CNR, the universities, the principal scientific society *Accademia Lincei* (page 117), the astronomical observatories (page 118), INFN, ENEA, and the ministerial office for research and technology (page 122).

Then follow three portraits: of a unique private research institute, the Mario Negri institute, just to show that success is possible in Italy (page 124); the International Institute for Genetics and Biophysics in Naples, in which success and failure combine (page 125); and the Zoological Station of Naples (page 127). □

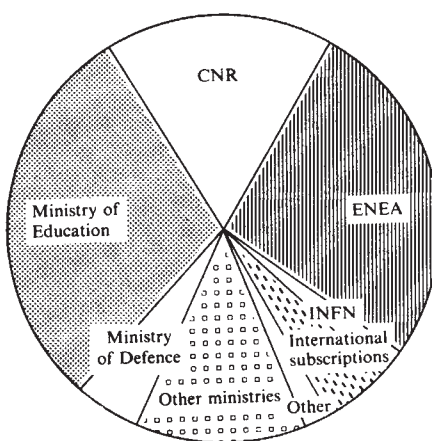
Counting the scientific lira

A question of imbalance

ITALY spends 1–1.2 per cent of its gross national product on research and development (depending on whom you ask), and in 1979 ranked sixteenth — behind even Iceland — in per capita spending (according to recent figures from the Organization of Economic Cooperation and Development). Spending appears to have been rising in recent years, but to have been choked back in 1982 because of national economic problems; but it can be difficult to keep track of where the money goes in Italy. It is not so much that the lira is a small quantity and that sums begin to sound astronomical (there are around 2,000 lira to the pound) but that apparently firm figures turn out to be meaningless.

Take the research budget of the ministry of education (MPI), for example. According to the annual report of the national research council (CNR), which reviews the whole of science in Italy at great length (the 1982 report is more than five inches thick) but little depth, the MPI budget leapt up 74 per cent in 1981. In fact many research budgets apparently soared that year, according to figures in the report: the budget for the ministry of defence shot up 300 per cent, other ministries 69 per cent. The CNR itself saw a 38 per cent rise, and the net public spending on research outside industry rose 63 per cent.

Inflation of around 18–19 per cent is not enough to account for the rises: even allowing for inflation, public research spending still seemed to rise by half in 1981.



THE main sources of finance for research and development in 1982 were the Ministry of Education (mainly supporting basic research in the universities), ENEA (the nuclear energy agency) and CNR (the national research council). INFN supports high-energy physics. International subscriptions include those for the European Centre for Nuclear Research (CERN), the European Space Agency (ESA) and the European Southern Observatory (ESO). Defence spending is low. □

However, nobody talks of these massive increases. What happened to them? Voted by parliament, published by CNR, officially available, most of these rises were in fact never spent. "We have great trouble spending money in Italy", says Dr Giorgio Sirilli, who prepares the figures in the CNR report each year, and whose work is respected by the Organization for Economic Cooperation and Development (OECD) in Paris. (The OECD science indicators unit prepares thorough international comparisons of science spending: the latest, taking data up to 1979, is approaching publication.)

One must distinguish, says Sirilli, between "finance" and "spending". In a tussle which seems to put the cart before the horse, the sources of finance (the ministries) are always trying to push more money onto the spenders (like the CNR and the universities) than the spenders can spend. Thus the "finance" figures, which appear in the parliamentary budgets and are easily available are always what might be called politically inflated: they represent the political ambitions of the ministers rather than reality, and are agreed by parliament with this comfortably in mind. True spending, on the other hand, can be limited in any year by controls imposed later by the treasury ministry, or by bureaucratic hurdles.

As a further complication, the general recession has struck Italy only very recently, much later than other countries; so monies promised are not always turning into monies given. A few rough calculations based on figures in recent CNR reports show that public *spending* (the real thing) on research and development in Italy in 1978 was 97 per cent of *finance* (the imaginary one), giving a 'political inflation' that year of only 3 per cent.

But the next year (1979) political inflation was higher: 8 per cent. In 1980, the last year for which true spending figures are available, it reached 17 per cent, and since the greater (though still unknown) part of the enormous increases in finance in 1981 were not spent, political inflation probably went through the roof that year.

As for 1983, the minister of research claims that Italian research and development finance will reach 1.2 per cent of the gross national product. But this is *finance*: the true amount spent is likely to remain under 1 per cent, just a shade higher than the 0.7–0.9 per cent it has hovered around since 1970, Sirilli believes. Even so, this indicates that research spending is rising in Italy, even if not as fast as the minister would like; and the finance figures also indicate an increasing political will to give support to science, something that is not at all negligible in a country run by a volatile coalition of five parties.

Biomedical sciences have fared particularly well, funding having *tripled* over the past five years according to some estimates. And in some institutes spending on astronomy has quadrupled.

Also, industrial research spending has been rising steadily in real terms, and in industry, differences between "finance" and "spending" are smaller than in the non-industrial sector and decreasing: 5 per cent in 1978, 2 per cent in 1979, 1 per cent in 1980.

But why is non-industrial research money not spent? In practice, there are simple obstacles like these: a research director of a CNR laboratory, for example, must spend the whole of his capital equipment budget in the year of allocation — anything unspent is lost, and an overspent budget would not be honoured. Getting the right forms through the right offices by 31 December can be impossible. Active directors like Professor Lucio Luzzatto, the haematologist, who directed CNR's largest laboratory, the international institute for genetics and biophysics at Naples, from 1974 to 1981 — can claim to have spent all their allocated budgets "but sometimes with difficulty". Others may not do so well. And at the level of CNR

itself, delays in getting planning permission for a new laboratory, or getting buildings built can frequently mean the loss — for that year — of the cash previously made available. Or the budget may indicate new staff: but other regulations forbid, or slow, their appointment.

As for the 74 per cent rise at the ministry of education — it is explained by the introduction of a major new fund created specifically for research in universities, as part of a new law for the universities which passed through parliament in 1980. The gross fund, said parliament, would be 91,000 million lira (around £43 million) in 1980, 141,000 million lira in 1981 and 191,000 million lira in 1982. But in practice much less than this has been spent, because the now more autonomous universities are still working out how to apply the terms of the 1980 law. "Without structures you cannot spend the money" says Sirilli.

And 303 per cent for the ministry of defence? An anomaly. "A new man was appointed at the ministry to give us the figures," said Sirilli. This man apparently decided that one category of spending that had been classified secret, and previously unreported, would now no longer be so. In fact spending was roughly constant. □

government nominees.

Thus CNR is said to be 'the universities' bank' (until three years ago, it handed out 90 per cent of university research funds) and 'in the hands of the barons'. (*Barone* is a common term of distaste for the professors — and a grudging admittance of their power. In translation it means both 'baron' and 'rogue'.) The barons, it is said, have used CNR to strengthen their castles in the universities.

However, since the widespread university troubles of 1968 the council has considerably increased the number of its own institutes, which gave CNR an identity and structure apart from the universities. It did this largely in response to the demands of university researchers, who found that they could not work in the now overcrowded and polemical universities, but then the remaining professors saw that CNR might go its own way, leaving the universities stranded and possibly underfunded. So the universities have jealously hung on to their power within the CNR committee structure.

But this great university power has left CNR riddled with anomalies. For example, despite the fact that the CNR committees control the council's own institutes, the CNR scientists and laboratory directors have almost no place on them. Of the 140 seats, 96 go of right to the universities, and only 20 are available to 'non-university researchers'. CNR researchers must find their place within that 20, alongside representation from other bodies. They are lucky to find one seat on each committee.

Perhaps it is not surprising that there are over 500 cases of financial irregularities, not all of them minor, pending against CNR. Partly in response to these obscure 'scandals', some involving land and property deals, the CNR budget was actually reduced by 7 per cent in 1982, in current lira (say 25 per cent down in real terms), compared with 1981. It seems a dramatic collapse, but CNR can cope with it, says a government spokesman, because of the slack previously taken up by the 'scandalous' spending. But Quagliariello offers another explanation for the cuts suffered by CNR in 1982: the recession.

Whatever the causes of the cuts, some institutes of CNR have been badly affected. The research institute of astrophysics in

The national research council

In search of an identity

THE Consiglio Nazionale delle Ricerche (CNR) is the principal research council in Italy, covering every field except high-energy physics, nuclear physics and ground-based astronomy, which are covered by INFN, ENEA and the observatories (see below). There is in Italy no medical research council, no agricultural research council, no social science research council: CNR does it all. Its budget for 1982 was 470,503 million lira (around £225 million), of the same order as the Science and Engineering Research Council (SERC) in Britain. CNR has 2,100 researchers of its own, has 2,300 technicians and other staff, runs its own research institutes, and makes a major contribution to university research (where there are 17,000 researchers) through grants and other means, such as 'study centres' and 'research areas' based in universities. But an increasing proportion — now around a quarter — of CNR's funds now goes to specific applied research programmes, the *progetti finalizzati*, and there is now a strong political will to make CNR the tool of government technological policy.

CNR manages — or fails to manage — this large budget and research community with a tiny central staff: a president (Professor Ernesto Quagliariello, a Neapolitan now nearing the end of eight years in the job), an administrative director-general, three central directors, and 25 graduate 'managers', a total of 30 people. The inevitable specialist advice (on

the distribution of grants etc.) and greatest influence lie with 140 scientists, mostly university professors, in 11 committees covering topics from jurisprudence to molecular biology. The committees rarely use referees. Compare SERC's more than 100 central staff, and 1,000 members of advisory committees, which cover a much narrower range of subjects. From one point of view, CNR is concentrated in too few hands; from another, it is under-managed.

The real power in CNR now lies with these university professors. They make up the vast majority of the 140 scientific advisers in the advisory committees, and the presidents of these committees, again usually professors, form the majority in the decision-making President's Council. This council is composed of the president of CNR, the director-general, the presidents of eleven advisory committees, and two



CNR president Quagliariello (left) and reforming "baron" Donato



Union man Spanedda (left) and university man (and Senator) Faedo

Rome, for example, a strong institute selected by the US National Aeronautics and Space Administration to be a European information centre for Solar System imaging, suffered a 20 per cent cut in real terms.

Everyone, including Quagliariello, now talks of a "reform" of CNR. It certainly needs it, although not because of financial scandals, which really are quite commonplace. There are two over-riding problems: the relationship between CNR and the universities (which dominate it), and the existence of a legal straitjacket which has been choking the appointment and careers of CNR's own researchers since 1975.

Professor Luigi Donato of the University of Pisa has been closely concerned with reform of CNR, serving on a committee advising the minister of research, Pier Luigi Romita, on exactly that question. Donato is a 'baron', in the sense that he is a well-established university professor. But he is not a baron of the roguish kind, according to local trade union opinion: Donato, they say, uses the devilish system to good ends. He is also director of the CNR institute of clinical medicine, 100 researchers working mainly on the heart and lung, but covering topics ranging from cell biology to the study of materials for an artificial heart and from patient care to epidemiology.

"The committee I'm serving in is the fifteenth [to attempt a reform of CNR]", he says. "But there is a large consensus now." Donato was hopeful that the reform could be voted through parliament by the end of the year, if the now out-going government had survived that long. The reform has taken so long partly because of the changes of government, but also because of opposition from the universities, which, says Donato, did not want to see a CNR structure "building up an independent strength".

The hope now for change stems largely from the 1980 law for the universities, which gave the ministry of education much increased finance for supporting university research, partly through a new, university-controlled structure, the national council for the universities. This increased university research finance, reducing the perceived importance of CNR and thus the resistance to reform.

The other great problem has been negligible recruitment and promotion in CNR, while 1981-82 saw the first new posts

— and plenty of them — in the universities since 1972. CNR has been almost unable to recruit, or promote on merit (rather than mere age) since 1975, when it fell, apparently willingly but under pressure from the trade unions, into a legal structure called the *parastato* (para-state) structure. INFN (high-energy physics) and CNEN (nuclear energy) shared CNR's fate. (CNEN escaped *parastato* status last year, when it also changed its title to ENEA; CNR and INFN are still struggling in the quicksand.)

Professor Lucio Luzzatto, who was director of the international institute for genetics and biophysics in Naples, CNR's largest institute, from 1974 to 1981, voiced his complaints about this in the official Communist Party newspaper, *l'Unità*, after the paper had announced that CNR was an "impregnable fortress" because of the entrance examinations it imposes on new recruits. CNR is an impregnable fortress, he said, but not for the reasons the paper stated.

In the first place, recruitment was blocked in 1975 pending detailed regulations that CNR was to work out on the form of recruitment, the wording of advertisements, and so on. The first of these regulations was not ready until 1978 and then there was further haggling over how many posts would be available, and how many there would be for each laboratory. By the time this was over, there were many people inside and outside CNR who had taken degrees and needed to be upgraded to researcher (a good thing, in Luzzatto's mind), but together with a general political requirement to recruit for the remaining posts first from the unemployed, the result was that the first effective external recruitment began in 1982. There had been seven years in which CNR was almost completely closed to outsiders.

"In most countries," wrote Luzzatto "it's calculated that research recruitment needs to be about 4-5 per cent per year; at present, they have to be content with about half. In any case, recruitment is planned to be steady." But in Italy, he said, the system seems to be one of fortress under siege, where "at unpredictable intervals the drawbridge is lowered and those who happen to be at the front are let in". Then the drawbridge is raised again.

The universities had also been closed, for different reasons (see page 116), and for

longer (10 years), but they opened first, more effectively, and at all levels. Also, for the moment, university salaries are higher.

In the past two years, more than a quarter of CNR researchers and technicians have moved to the universities, reversing the trend of the early 1970s. (The CNR establishment was 6,100 at the beginning of 1981; now it is 4,400.)

Another problem with CNR jobs is that recruitment is only possible from the bottom. Professor Donato was part of the CNR delegation in 1973 negotiating with the unions over the *parastato* question: "I heard the theory of the cylinder," he said. A man or woman comes in at an early age; and rises by seniority. There is no recruitment at higher levels from the outside as this would be unfair to those inside. The union delegations were mostly composed of technicians concerned about security, says Donato; but in 1975 the unions won, and CNR and other institutions fell into one class with entirely bureaucratic organizations where promotion is by seniority. "So the only way to progress in CNR is to age — to survive" says Donato. "The main complaint of a university man is that there is no career in CNR: there is no mechanism for selection."

In 1975, however, *parastato* status had its attractions for researchers: it offered security. There were only 33 permanent posts in the organization and its laboratories before *parastato*, the same number as when the CNR was founded in 1923. The rest of the staff had to be content with five-year contracts.

But three years ago, CNR researchers started a strong movement to modify the

No room at the inn

NON-Italians have been getting a raw deal at CNR — so much so that they have formed a pressure group and are taking their case to law. Despite living in the country and working for CNR for upwards of a decade, some even taking Italian nationality, these 18 scientists and two secretaries cannot win permanent jobs with CNR, while Italians in similar grades are so appointed. The 20 are all European, and include prominent scientists, such as a Spaniard (now Italian) who runs a major part of a CNR 'finalized project' on informatics, and a German who helped establish molecular biology in Naples (Werner Schreil, see page 126).

The CNR position is both contrary to natural justice and a violation of the Treaty of Rome (which established the European Community and calls for freedom of employment within the community) says one of the 20, plant virologist Robert Milne of Turin. It seems short-sighted: foreign scientists could be useful to Italy. In Rome, CNR puts the blame on its *parastato* status, which limits its freedom to employ foreigners. But this does not explain the reluctance to appoint the newly-naturalized Italians. . . □

parastato status "and they are succeeding" says Donato. The unions are "kind of accepting" that things must be changed. According to Dr Luigi Spanedda, a prominent member of UIL, the socialist union, for example, "promotion should be by scientific merit, but the present structure doesn't allow it". But this was a personal view, he stressed: others in the unions "think the same but dare not say".

The unions may make trouble on other grounds, however: the present reform (under discussion by Donato's committee) is much weaker than one proposed eight years ago by the Christian Democrats, the Communist Party and the Socialist Party, says Spanedda. The professors will still have the majority in the committees and the unions will certainly oppose it for not sufficiently reducing the power of the barons.

According to Donato, the new CNR structure would offer a three-level career to researchers, parallel to the new one for university professors (set up by the 1980 law for the universities). Passage between the levels would be by merit, and there would be an open contest for recruitment or promotion at each level. "It would be an open system — that's very important."

The names of the jobs would be changed, too. CNR researchers are now called 'technical collaborators', a term the researchers consider offensive (it has the same root as the word for 'maid'). This "nonsense" would be changed to 'researcher' Donato says. The career of technicians would also be separated from the career of researchers: and the unions have agreed to that.

As for the committees, now so dominated by the universities, university membership would be reduced to 55 per cent (too high for Spanedda, as it still represents a majority) and CNR researchers given their first explicit places: 25 per cent would be reserved for them. The remaining 20 per cent would be for outside experts. It is correct to give the universities the majority here, Donato feels, "particularly since their new role will be only advisory". Moreover it would not be possible to get anything less through parliament, "which is heavily loaded with professors".

At CNR headquarters (in Rome) there would also be changes. The President's Council would become advisory and there would be a new body, a 'board of directors' to take decisions. This would be composed of around 11 people, of which 4 would be university members elected from the advisory committees, 2 or 3 from CNR (so that these first two groups formed the majority), and the rest nominated by the government on the basis of expertise, particularly from industry.

The danger of this system is that it might become heavily politically loaded, says Donato (political 'clientism' being very strong in Italy — the CNR is traditionally presided over by a Christian Democrat, for example) but "it's worth trying". □

North and south

Can the regions run biology?

FINANCE for biomedical research has tripled in Italy over the past five years, according to Professor Rossi-Bernardi of Milan, who keeps an eye on such things. It now approaches 100,000 million lira (£48 million), but the increase has come at the expense of a collapse in coordination which threatens coherence.

The money is coming in three parts: from the national research council (CNR), from the national council for universities, and — in principle — from the regional health authorities. But apart from two regions, Lombardy (around Milan) and Tuscany (Pisa), the authorities have failed to set up the local legislation and management structures which are required by national law for the money to be spent, although the total sums involved are greater than the previous whole budget of CNR for biology and medicine!

The national law dates from the setting up of the Italian national health service, on something between British and Swedish lines, as recently as 1978. (Before that, medicine was private, paid for mainly by health insurance provided by a person's employer.) Within the law was a provision for the 20 regional governments to spend a fixed fraction (0.35 per cent) of their health service budgets on medical research. Five years later, little of the money is being spent.

So is there a danger of this promised money being lost to research? Yes, there is, says Professor Luigi Donato of the University of Pisa. "All new money is dangerous money: it can go anywhere."

And when and if the funds are released the spending will have to be coordinated so that a research project turned down by a larger body, like CNR, the national council for universities or the European Community would not automatically be

picked up at regional level. Donato is executive vice-chairman of an inter-ministerial committee, linking the ministries of education, health and research, which is looking into this coordination problem "and I can tell you it's a very tough kind of problem". The committee believes that the regional fund should be spent on 'health service research', problems related to health care delivery. Or research already strong and with particular importance to medicine or industry in the region could be supported. "We will end up with coordination between the university council and CNR", but the regions are constitutionally very strong and independent and not to be ordered about.

Moreover, although the regions are required to set up a management structure for distributing and controlling the funds, they are actually free to set up almost any kind of structure they like. This is "absolute nonsense" says Donato. In Tuscany (Donato's region, where he has influence) the authority has established an advisory committee composed of two CNR members, three university members and so on, which will be in charge of assessment of projects. The region has 5,000 million lira to spend (£2.4 million) each year "which is really a lot of money".

The two great risks in the regional funds are wastage, and the support of nonsense research. However, Donato is not so worried about the big regions, like Lombardy, which is the size of a country like Belgium and has three universities and plenty of excellent researchers. "It is very unlikely that mistakes will be made there." The problem is in the south, where "the big danger is that they will try to develop local experts to fill their committees" when those experts do not in fact exist. □



Cave houses in southern Italy, abandoned only 30 years ago

Italy's 'top' scientists

Four in the top one thousand

AN interesting measure — or at least indication — of the productivity of Italian scientists can be found in *Science Citation Index*: the list of 1,000 contemporary most-cited scientists, published in October 1981. The list took an Institute of Scientific Information (ISI) team two years to prepare, and required 750 hours of big computer time: it ranked the number of times a scientist's papers in 3,000 major journals had been cited in the 14-year period 1965–78. The study covered five million articles and 67 million references. Top of the 1,000 was cancer immunologist Robert Good, Chairman of the Sloan Kettering in New York, with 17,679 references. Scientist number 1,000 on the list scored 2,436 references.

The list contained 41 Nobel Prize winners, but only four Italian scientists working in Italy, who were:

- Eraldo Antonini, biochemist, laboratory of molecular biology, University of Camerino (in north central Italy), born 1931: 3,127 citations;
- Enrico Clementi, theoretical chemist, 'G. Donegani' research institute, Novara (near Milan), born 1931: 4,001 citations;
- Silvio Garattini, pharmacologist, Mario Negri institute for pharmacological research, Milan, born 1928: 2,833 citations;
- Giorgio Giacomelli, physicist, University of Bologna, born 1931: 2,483 citations.

At least six names on the ISI list of 1,000 were also Italians, but working in the United States; evidence of a brain drain.

One of the four remaining in Italy, Professor Garattini, is interviewed later in this issue (p.124), as director of the independent Mario Negri institute.

In a separate study, professor Rossi-Bernardi of Milan finds that three-quarters of the biological papers published by Italians in 1981 were written by university researchers. There is little inclination among biologists (other than the agriculturalists) to use Italian journals: only 15 per cent of the Italian biological papers are published in Italy. Only 8 per cent of the papers were written by researchers from the independent institutes of CNR, roughly in proportion to numbers of researchers, but over 50 per cent of all researchers were receiving their prime funding from CNR, so in 1981 it was still acting as the universities' bank, despite the entering into operation of a new fund for the universities at the ministry of education.

The figures show that the number of Italian biomedical papers in quality journals has risen by a quarter in the past two years (which follows a rise in funding), and indicate that the volume of papers produced by Italy is not far short of that produced by France or Germany, where the research spending is much higher. □

Industry and CNR

Applied research — who benefits?

THE role of the Consiglio Nazionale delle Ricerche (CNR) in relation to industry is ambivalent. CNR houses many important basic research institutes — yet it is now widely considered that CNR should become the "zip" (ex-research-minister Giancarlo Tesini's term) linking the universities with industry. The feeling is strengthened by the fact that the universities now have their own direct funding system from the ministry of education.

According to professor Ernesto Quagliariello, president of CNR, the existence of this funding frees CNR from its previous main business, the universities: "according to me, CNR must now concentrate on applied research . . . I do not say that CNR support for the universities should be blocked, but it should continue on a new basis, as a bridge between universities and industry". As for the basic research conducted in CNR, he says: "CNR has 300 laboratories. I cannot exclude basic research in all of them."

But Professor Francesco Blasi, director of the largest of all the CNR institutes, the international institute for genetics and biophysics in Naples — which does basic research in molecular genetics (see opposite) — says: "Eventually, CNR will become an applied research council . . . and when it happens, Italian science is going to drop even further. The real results, the new things come mostly from CNR and not the universities."

In fact CNR has been acting partly as a "zip" or a "bridge" since 1976 in the so-called *progetti finalizzati* or 'finalized projects', nominally applied, five-year research projects which accounted for 25 per cent of the 1982 budget of CNR. The *progetti finalizzati* are a product of the oil crisis, and a political realization that Italy had to do something about research and development. The bulk of a first batch of projects ended in 1981–82, and a new batch was approved last year.

Strictly, CNR only manages the projects, but they were first established by a committee (now dissolved) formed from the university-dominated committees of CNR, and the links have remained intimate.

So there is already considerable experience of how zips between Italian universities and industry might work. Opinions differ widely, however, on the interpretation of that experience. What seems to emerge is that the finalized projects have been useful, but not necessarily in the way intended; that they have been widely abused, and rarely applied; and that a second set of projects approved last year, might do better. (The word 'finalized', is not really a good translation from the corresponding Italian. In Italian the word means 'directed to some end or purpose' whereas the

English word tends to mean simply 'finished'. So *progetti finalizzati* are really 'goal-oriented projects', but the phrase is a mouthful and we stick to 'finalized projects' here.)

So what of opinions of these projects, of which there were 23 by 1981, divided into some 10,000 individual grants in fields such as health, energy and agriculture, having accounted between them for 480,000 million lira (around £230 million) since 1976?

Take the views of Giancarlo Tesini, the previous research minister, for example. A year ago he reported to CIPE (the economic planning committee) somewhat predictably that his office did not have the means to oversee the projects properly (this is an oft-repeated and well-founded complaint of the minister), so there was a lack of control. Also, said Tesini, the projects had been defined by scientific discipline, rather than by the goals which were supposed to be their object. Thus, he said, this led to "an excessive freedom of choice in the management of the projects". This freedom was left to the project committees, which were (and are) composed mostly of university professors.

Tesini also complained that potential users of the results were too little involved in the original project definition; that even the scientific objectives were ill-defined; and that what administrative support there had been was bureaucratic. Most important, there was no means established of using the results. For example, in the agricultural projects, said Tesini, it was "urgent to create an organic system for distributing the results of the finalized projects to the central and regional administrations and to the farmers". In other words, so far the results had remained with the universities.

Another complainant, from the shopfloor so-to-speak, is Dr Luigi Spanedda, a union man from Pisa. "CNR committees invented them," he says "to convince politicians to continue funding CNR". (In fact, after the oil crisis in 1973, which hit Italy hard, it being the highest industrialized importer of oil after Japan, biology funding — for example — fell by half; Spanedda's thesis is that the finalized projects were the universities' reaction, rather than the politicians' initiative.)

"For example take the energy project," says Spanedda. There was no national energy plan at the time. A few people on the committees asked colleagues to propose ideas; this led to programmes designed to support particular groups rather than designed to reach a particular technological end. It was a *mecanismo perverso*, a 'corrupt mechanism'. Another example was the finalized project on containers, apparently very technological and applied:

"the scientists say it was excellent, but the market is not interested . . . It was re-inventing the wheel."

"In a country like Britain," Spanedda emphasizes "applied projects are contracted by different groups from the customers. This applied/basic division does not exist in Italy."

However, there are positive views of the finalized projects to be heard, even in industry. Some 40 per cent of the non-administrative funds of the finalized projects went to universities, but industry got 30 per cent. Alfa Romeo used one to develop an electronic modular engine, which switches from four- to three-cylinder operation when acceleration is not needed and when it would save petrol (see below).

Also, the finalized projects enabled groups which were otherwise isolated to get together. Dr Constanzo Panichi, for example, a CNR researcher who heads a sub-project of the energy project dealing with geothermal energy, which started in 1977, says that the project "created a community", particularly in exploration. The project utilized geothermal wells, isotope geochemistry, and a few special wells (10–15 holes of around 100 metres depth) to measure geothermal heat flow, and worked on the design of a low temperature (80–130° C) two-fluid heat engine.

Panichi's project involved 550 man-years of work, was 70 per cent basic and 30 per cent applied, Panichi says. The number of people working in the field rose, and a major result was a map of the subsurface temperature of all Italy at various depths: the first thorough geothermal map of the country. Now a second project is under way, working on the evaluation of resources, on determining the permeability of deep rock from the surface (important in reducing drilling costs, but difficult), the development of heat exchangers, and practical applications: where special problems arise. There is no legal definition of who owns the heat, and the local authorities are not interested in funding developments until they know. What's needed, says Panichi, is a national law defining rights and responsibilities over renewable energy sources. But these cannot be said to be difficulties with the finalized project itself: it was a success.

Another successful project, even admitted by Spanedda, was "the first self-synchronizing heart pacemaker in the world", using sensors and a microprocessor, which was developed at the CNR institute for clinical medicine in Pisa by Professor Luigi Donato's group. "He got industry interested," says Spanedda "so the market gave its objective view by participation".

And in Turin, Professor Osvaldo Lovisolo, director of the CNR institute for applied plant virology, says the finalized projects were good because they obliged people in different organizations to work together.

Now a new series of finalized projects

'Finalized project' helps Alfa save fuel

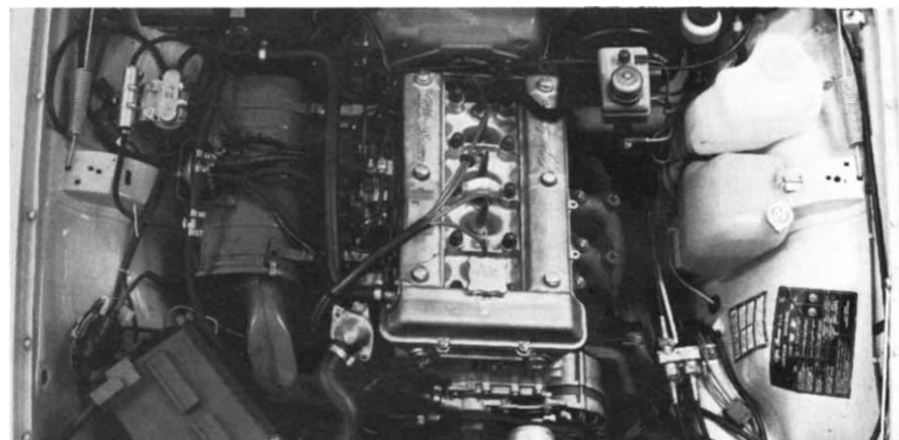
DESPITE some complaints that CNR's *progetti finalizzati* (see below) benefited basic science more than their object — applied research — at least one of them bore the fruit intended. This was the energy project, a sub-programme of which has helped Alfa Romeo, the state-owned car maker, design a fuel-saving engine. (Funding from Brussels helped, too.) Some say it's crazy that Alfa, renowned for its up-market fast cars, should think about fuel saving, but — says Alfa — by starting from racing engines the company has developed an approach which will ultimately give it a world lead in the technology.

A few years ago Alfa returned to racing (after a gap of a couple of decades) to stretch its technologies rather than gain prestige. Its engineers learned to treat each cylinder of the engine as a separate unit. A racing engine runs at very high temperatures, cooling is not uniform, and cylinders run cooler and suffer different stresses from the central ones: each cylinder needs different treatment. The same is true of an ordinary mass-production engine, though the effects are smaller.

Thus when Alfa Romeo turned to fuel-saving in its production-line cars, the company approached the problem cylinder-by-cylinder. At the same time, Alfa developed its version of electronic ignition and control of the engine — and the result was the 'modular electronic engine', which switches from four- to two-cylinder operation when the power requirements are low.

In a test of the engine in ten Milan taxis in 1980–81, Alfa found that mere electronic management of a normal engine, without the two-four cylinder switch, could save 12 per cent on normal fuel use; and that with the two-four switch savings rose to 24 per cent. Of the 21 taxi drivers who drove a modular taxi "long enough" more than half found it "good", 40 per cent found it "excellent", and only one "fair".

The switch from two to four cylinder operation can make the car lurch slightly — which can be bothersome in traffic — but Alfa engineers hope to solve this with an improvement in the software which judges exactly when to make the change. Some 800 modular Alfas are already going on restricted sale, but full marketing will have to wait until 1985, says the company. □



has been approved, although — perhaps taking account of Tesini's and other criticisms, and perhaps reacting to Italy's enormous budget deficit of around 30 per cent of gross national product — CIPE has not allocated as much money as before: 282,000 million lira (£134 million) for 1982–86.

The old projects fell under the broad headings: food and agriculture (44,000 million lira, or £21 million, from 1976 to 1982); human health (63,000 million, or £30 million); the environment (82,000 million lira, or £39 million); advanced technology (98,000 million lira, or £47 million); energy (71,000 million, or £34 million); and others amounting to 124,000 lira (£59 million). Industry became by far most involved in just four projects or sub-projects: those on energy, informatics (within advanced technology), biomedical technology (within human health) and fine chemistry (within advanced technology).

The great innovation in the new projects is the tremendous emphasis on medicine and biology, and on agriculture: both show an absolute rise in funding, with biomedical topics rising to over 41 per cent of the total (13 per cent before) and agriculture to 43 per cent (less than 10 per cent before).

The new projects agreed by CIPE are entitled: the structure and evolution of the Italian economy (a social sciences project) (13,000 million lira, or £6 million, from 1982 to 1987); mechanical technology (31,000 million lira, or £15 million); raising agricultural productivity (122,000 million lira, or £58 million); preventative medicine and rehabilitation (67,000 million lira, or £32 million); infectious diseases (16,000 million lira, or £8 million); biomedical and health technology (14,000 million lira, or £7 million); and for the first time explicitly, genetic engineering (19,000 million lira, or £9 million). □

The universities

Getting back on the rails

THE 30 or more Italian universities (the number depends on the definition) now cater for a million students, with a teaching, research and assistant staff of 42,000. Entry, since 1968, is open to anyone who graduates from high school — or is over 25 years old. In some cities, like Messina, a tenth of the local population is enrolled at the university; but typically one third of those enrolled never attend a lecture. The University of Rome is the largest university in Europe, with 140,000 people on its lists. The University of Naples is not far behind with 100,000: first year biology in Naples caters for 1,200 students. Medicine and law are particularly crowded: at the University of Bologna, for example, 25 per cent of the students are studying medicine. Law hovers around 15 per cent. Student courses last between 4 and 6 years, and two-thirds of those enrolled do not make it to the end of the course. It is a very taxing and inefficient system for the lecturers. Somewhere in the *melée*, most of Italian science is done: 75 per cent of Italian biological papers quoted in *Current Contents*, for example, emanate from the universities.

These giants of higher education are the result of nearly two decades of fundamental strife, which left wounds only just beginning to be healed by a series of laws voted in 1980 which restructure university careers and finance. But still some say the medicine is not the cure.

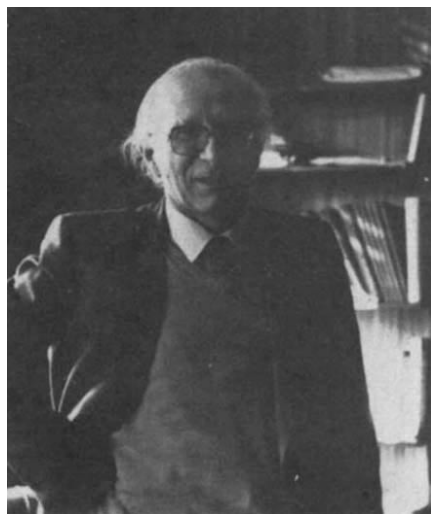
The problems date back to 1964, when the Germanic *privat-dozent* system of appointing professors was abolished. According to Professor Luigi Dadda, rector of the renowned Milan Polytechnico (which specializes in engineering and architecture, and played a key role in Italy's industrial revolution, educating people like Pirelli, for example) the system was "very serious", the first step in a professional career. "Before that you were an assistant, doing research for five to ten years — then you underwent an examination at national level on your publications and your teaching ability." This happened at the age of 30–35. "And it wasn't simply counting publications." The examiners analysed the content of the papers, "and there was real discussion". It was a goal for the assistants, and the forge of quality in the university. The system was abolished, says Dadda, because of a fuss over certain medical people who were using the consequent title of professor to practise medicine, and other factors.

But whatever the merits or demerits of the *privat-dozent* system, nothing was put in its place in 1964: after the abolition, there was simply no way left to appoint professors. "One of the mysteries of this country," says Dadda "is that something can be abolished without anything to

replace it". The laws of 1980 introduce, for the first time, a kind of PhD in Italy, a 3 or 4 year 'doctorate of research'; and the first of these are just entering university in 1983. Dadda sees these degrees — there will be around 2,000 a year — as the first substitute for the *privat dozent*. "But only now do we have a doctorate, 20 years later!"

The situation created by the decision of 1964 was made even worse in 1972, when every kind of turnover system in the universities, at all levels, was blocked, pending a new law which was expected the next year. This law did not appear until 1980, by which time there had been no new appointments of professors in Italian universities "for a full generation", says Dadda.

In the meantime the universities coped by "giving provisional professorships — it



Professor Dadda — mourning the 'privat-dozent' system

was very difficult to manage". Many good scientists left Italy after 1964. Then there was the student revolution of 1968, which went further in Italy than anywhere else, even France. The upheavals of 1968 led to the opening of the doors, and for those not politically inclined, too much polemic and too little research. Teaching suffered too: according to a physicist from the University of Bologna "it was like a return to the days of Galileo: you had to teach the best students at home", things being so confused in the universities themselves.

After 1968, many university researchers left to join CNR (Italy's national research council) and set up laboratories there. Nineteen sixty-eight had three major effects on research, according to a CNR astrophysicist: professors spent more time teaching; middling research was thus wiped out, but the best survived, as the motivated researchers managed somehow; some researchers moved to CNR, changing the nature of that institution; and some research-

ers emigrated.

But 1980 was a watershed. In just a few months some 15,000 new posts were created, outnumbering the 3,500 existing professors, and provisions were made for a total of 45,000 posts, even outnumbering the total number of people who were at that time, by hook or by crook, often without pay, actually teaching or helping at the university (42,186 in the academic year 1979–80). By now many of the 'new' posts have been filled from within this pre-existing staff.

According to Professor Luigi Donato of the University of Pisa, the 1980 laws "were like the laws you make after a war, after disasters, when you have refugees and orphans to look after. They are there, what can you do with them?" The effect "was merely a rationalization of the existing situation," says Professor Fernando Nicolo, engineer at the University of Rome. "This was the main spirit of the laws of 1980."

Although there is some provision for new appointments, the result has been to block the system with a generation of young people, and close it to younger ones, for a further ten years, says a professor at the University of Turin. "The appointments were not open to all people — only those already at the university." More than two-thirds of these people have taken up the offer, which required passage through examinations, but it is generally considered that anyone who wanted to stay could stay, and that selection was minimal. Even so, among those who stayed were many of high scientific standing, and they certainly deserved and needed the posts.

The 1980 system, which generally goes by the name 'law 382', as laws are usually identified by numbers in Italy, defines three levels of university staff: researcher (15,000 posts); assistant professor (15,000 posts); and full professor (also 15,000 posts). There are tests of competence between each level, but since the numbers appear to imply automatic promotion, it is difficult to see how the tests will take effect.

There are three important innovations in the law, however: the introduction of the 'PhD'; the provision of a new source of research money through the ministry of education, by-passing the old system in which CNR funded most university research; and the granting of greater autonomy to the universities.

The universities will be free to organize their own teaching, and their own forms of departmental structure; and in addition to the three levels of professor a university will be able to hire 'contract professors' for a year or two from abroad — at whatever salary seems appropriate to the university and satisfactory to the visitor. "These are very good aspects of the law," says Donato.

The real concern is for new posts: all around the country law 382 has led to a feeling of relief, but a concern about getting new blood. The fellowship system has

practically disappeared in Italy, as a kind of reaction to the inevitable permanence of any appointment lasting longer than one year — a right fought for by the trade unions and enshrined in law. But the main mechanism for testing new researchers is a part-time or limited-duration contract system, says Donato, "and we are badly lacking that". The only existing mechanisms are the new 'PhD' system, and a very few fellowships available in biomedical research through the new regional health funds.

As for the 'PhDs', the complaint is that there are too few, just 2,000 this first year, with probably another 2,000 for each of the next two years, leading to a steady population of 6,000 students. For the nominal 45,000 professors this is very few students, even if only half the number of professors is really seeking to do research; and the result has been that they have been fighting over them.

One senior ex-CNR biologist says that professors are getting together in groups of five or six to share a student. The students are not guaranteed a university post when they have finished — although given the small numbers, and the careful pre-selection of the students (in complete contrast to the free entry at undergraduate level), it is unlikely that they will not find a job in the university at the end of their studies. Indeed Luigi Dadda's automatic comparison of the 'PhD' with the *privat dozent* suggests that this will prove to be the case. The result will be that the groups of professors will 'claim' the post as theirs in rotation, and insert their own nominee, maintaining the old system of the grand professor and his acolytes which the 1968 revolution had attempted to overturn, the CNR man expects.

Thus according to Donato "the universities are rapidly digesting the innovative aspects of the law". And according to another cynical but progressive professor, the history of the university posts (and of CNR posts, which suffered a similar cycle but for apparently different reasons) "is ideal for the barons [professors]" for it enables them to exercise their *de facto* power within the apparent chaos. As for the sudden increase in the number of professorships, which appears to devalue the title of 'professor', it will not affect the true barons, this man says, for "they will find other titles for themselves within the new system, such as chairman of department or something".

Professors at the University of Rome, the biggest of the Italian universities, reject the charge: the style has changed, they say, laughing: "we no longer have post-graduate students working for us, we work with them". Professor Alberto Isidori, dean of engineering at the University of Rome, for example, is clearly uninterested in assembling baronial power: the great opportunity of law 382 for him is the opportunity of creating true interdisciplinary departments, avoiding the

The many lives of the 'lynx'

NOT many scientific societies can claim to have been established before the trial of Galileo Galilei — and to have had this father of science as a past member. But the *Accademia Lincei* (the academy of the lynx) in Rome has this honour.

The academy was founded by Prince Federico Cesi, then a young man of 18, in 1603. "We think it is the first scientific academy in the world" says its present president, geneticist Giuseppe Montalenti. The *Accademia Lincei* predates both Britain's Royal Society and the Académie Française by around sixty years; only the small Sicilian society, the Frederick II Academy of Palermo, claims an earlier origin.

However, while Prince Cesi chose the symbol of the lynx to emphasize sharp sightedness, he might have chosen it to recall the many lives of the cat. The *Accademia Lincei* has died — twice — for its belief in the independent truth of science. After a brief first existence it collapsed with the trial of Galileo (in 1633, exactly 350 years ago this year), and it died again under fascism when Mussolini demanded an oath of loyalty from its members and a say in their appointment.

After each 'death', the academy was reborn — although the period of limbo after Galileo was long, lasting until Quintino Sella, a champion of science after the unification of Italy, re-established the academy in 1875. Sella gave the academy the Corsini palace, a glorious building set in the botanical gardens of the University of Rome, a few hundred yards from the Vatican; and the academy instantly made

Charles Darwin one of its first new members, which was quite a brave move so close under the eye of the Roman Catholic Church.

Now, however, it suffers from a degree of ossification because of its fixed membership: by constitution it must have just 144 national members, 144 corresponding members (a junior class), and 144 foreign members. The average age of the members is around 75, and new members are appointed only on the death of the old one. Montalenti, himself a sprightly 78, wants to increase the membership, perhaps by introducing a class of 'emeritus' members, but such a constitutional change requires a majority of existing members to agree to it. "Some are old, some are sick, it's difficult to get them here" says Montalenti. And is the membership conservative? "Yes, very much so."

The effect is important: that the influence of the academy — as distinct from its enormous prestige — is much smaller than it would be if Italy's younger and more active researchers were members. Thus Italian scientists effectively lack a central voice. The academy rarely attempts to advise government, and when it does, Montalenti is somewhat despairing of the politicians' attention.

Montalenti's present interest: conservation areas. "We are destroying our environment in Italy" he says. There are strict laws, but they are ignored. The ministries listen to the academy, and arrange meetings about the matter. "But it's just *pour parler*" says Montalenti. □

wastage of resources (such as duplication of libraries) inherent in the present system of institutes, one attached to each professor, a system which also makes novel approaches to research difficult. "It's true that departments could become super-institutes," he says, but the temptation has to be avoided.

Also, says Isidori, although the law imposes a cylindrical demography, including — by implication — many mediocre researchers, a "spontaneous recognition of each other's abilities" within this cylinder can lead to the formation of effective groups. Those outside the groups will not absorb resources at the same rate, he hopes, so dissipation is not inevitable. "We must try to get the best out of it" he says.

Another problem, though, is finding effective technical support. According to figures published by CNR, there are six university researchers to every qualified technician. The CNR institutes do much better with a ratio of 1.6 researchers per technician: which is another reason why some scientists prefer to work in CNR. One full professor of biology at the University of Turin says he has no access to a secretary, no technicians "and I even have to

open the lecture room door to my students". But there is great diversity, some laboratories being relatively well established.

As for the new money from the ministry of education, it has gone like this (according to figures compiled by Professor Rossi-Bernardi of Milan). The total sum planned for distribution in 1980 was 91,000 million lira (£43 million); and that sum was in fact distributed. The next year the sum was to be 141,000 million lira (£67 million), but only 121,000 million lira (£58 million) was spent. In 1982 (last year) the sum was to be 191,000 million lira (£91 million) but the probable out-turn will be less even than 1980 at 70,000 million lira (£33 million). This is just the way things seem to go with money in Italy: unpredictably.

As for its distribution, which some have described as "sprinkling", 60 per cent goes straight to the senates of the universities, according to a simple formula depending on numbers of researchers; and 40 per cent is distributed by a system of 14 committees dubbed the 'national council for the universities' (CUN) which responds to grant applications according to research priorities 'of national interest'. So, for ex-

ample, in one year the University of Rome received 10,000 million lira (£5 million) as its share of the 60 per cent. It gave half to the various research schools of the university (there are 13) according to their broad needs, to provide a minimum necessary for research. The rest was distributed by internal competition, with decisions taken by a committee appointed by the senate. The bulk of the money from the ministry was thus distributed selectively, the university claims.

The other main source of research money for the universities is CNR; but the character of this support has changed in the past five years, becoming more goal-oriented through the 'finalized projects' (see page 114). "But they are not actually

finalized — they are just a different way of doing basic research", a University of Rome man admits. One complaint about them, from the university point of view, is that they were planned "outside the universities", by professors in the CNR committees who used them to feather their own nests rather than support university research in the broad. On this view, there is a difference between 'professors in committees at CNR' and 'the universities' — which is, of course, true for any research council in any country. As for professorial dominance of CNR: that's inevitable, the professors say, because the CNR research community (2,100 researchers) is too small to look after itself. The CNR researchers, however, beg to disagree.



Southern Italy may provide viewing conditions as good as at the Anglo-Australian Telescope, says Franco Pacini

Astronomical Italy

Funds quadruple but $f = mv$

GALILEO'S ghost must have been wringing his ghostly hands! At the Arcetri Observatory near Florence, just a few hundred yards from the Villa Gioiello, where Galileo died, astronomer Franco Pacini was explaining that in Italy, $f = mv$. (Galileo arguably showed that for the weight of a body, $f = mg$, where g is the acceleration due to gravity. Newton extended this to $f = ma$, where a is a general acceleration.)

But Galileo may rest in peace. Pacini, who directs the Arcetri observatory, and is himself a powerful force in Italian astronomy, is talking about Italian science politics. "If you stop pushing, things stop . . . $f = ma$ means if you get something going, if it's decided, it goes ahead. Here, instead, $f = mv$. Something gets decided, but then unless you keep going to the ministry, unless you keep pushing it, it doesn't go through . . . This destroys people — spending 70 per cent of their time carrying through things that are already agreed."

"Another problem is that sometimes you get money, and sometimes you get people [positions]; but the two are not in phase." Nevertheless, "you can achieve results. It just requires much more effort than in other places".

And Italian astronomy has seen big results since 1978, though there are clouds on the horizon now, as the recession is biting in Italy somewhat later than in other countries. Arcetri, an important astronomical centre in Italy, has seen its funds *quadruple* since 1978: and while Arcetri may have been especially privileged, Italian astronomy in general has certainly seen a brief "period of grace".

In this period, the astronomers took their opportunities, developing important new national facilities which at last give them the opportunity of being more than mere guest observers at other institutions. Arcetri has built and runs, since last November, an infrared telescope on Gornergrat mountain near Zermat in

Switzerland ("the best in continental Europe" says Pacini); the first of two 32-m dishes for very-long-baseline interferometry is ready in Bologna, with the other to be in place in Sicily next year; a computer network for data analysis (similar to the British Starlink) is already working; the design of a hard X-ray satellite "for the end of the decade" is underway; also there are firm plans for a national 3.5-m optical telescope. Italy last year also joined the European Southern Observatory, ESO, whose telescopes are in Chile. Also 100 new posts have been promised: a 25 per cent increase in the number of astronomers in Italy. Altogether an impressive catalogue for just five years of grace.

The ten Italian 'observatories' — which dominate ground-based astronomy, and still do some observing but now act mostly as home bases — have also benefited from institutional reform. They are very old institutions, independent of other scientific bodies but closest to the universities: observatory directors are usually professors of astronomy at the nearest point of learning. Some have been hidebound and reactionary; others, like Arcetri, have been more forward-looking. But now their isolation will be much reduced, in a reform which followed the reform of the universities. Directors previously had the job for life; but now the term is reduced to three years (renewable). Careers for astronomers in observatories which used to be based on seniority alone, are now parallel to those in the universities, with three tiers and promotion by merit. Exchange of staff between university and observatories is eased, and the observatories are thus 'opened up'.

"We have essentially been established as a kind of institute of the universities" says Pacini "and with access to the same funds [from the ministry of education]".

The law for the observatories also com-

mits Italy to building the proposed 3.5-m optical telescope — although it does not yet provide the funds. "There are rumours that the ministry already has the money in its pocket — but they haven't written us a letter yet to say we can have it." The telescope has been the astronomers' dream since the 1950s — but now there are "good prospects" that it will be realized, says Pacini. To save money, it would be near-identical to ESO's new telescope, now in the design phase, using "new technologies: a tin mirror, rotatable dome etc. . . ." It might go to the Canary Islands — though these are looking increasingly expensive — or to the south of Italy, where observing conditions match those of the Anglo-Australian Telescope, Pacini claims.

However, as the economic clouds gather, Pacini is already getting indications that money is getting harder to find. "Our infrared telescope represents an investment of 2,000 million lira (1 million) but CNR told us we might not be able to find 20 million lira to run it . . . And if that had been so we should have raised hell!". Which, clearly, would have brought the period of grace to an abrupt end. However, on 28 April the money came through, two days before closure was inevitable. Just another example of Italian troubles, says Pacini.

Italy in space

WHICH country in Europe spends most on space technologies as a proportion of its total research and development effort? France, you say? No, the answer is Italy, and by a long way. Some 9.5 per cent of Italian research spending was devoted to space in 1979, according to European Community statistics. France weighed in at 4.5 per cent.

Behind the increase is a national space plan, designed to bring Italian industry into the space age. The plan got under way in 1979–80. Present key projects: ITALSAT, a telecommunications satellite for launch 1987; IRIS, an upper stage for the US space shuttle; TSS, a 'tethered satellite' that the shuttle will tow through the upper atmosphere; and with the Dutch, SAX, an X-ray astronomy satellite.

High-energy physics

The Enrico Fermi legacy

ONE scientific subject always seems to get a fair amount of political and, lately, an enormous amount of public attention in Italy: high-energy physics — or elementary particle physics, to use a parlance that Enrico Fermi would have better recognized.

"Fermi, Marconi — it takes a long time for the effects of those volcanoes to die away" says Umberto Vattani, the Italian chairman of the council of the European Nuclear Research Centre (CERN). The tradition of Fermi in Italy has been pursued by INFN, the national institute for nuclear physics, which has an important peculiarity among Italian scientific institutions: its president is not named by the government, but elected by the physicists.

This, combined with the memory of Fermi, gives the president of INFN an unusual authority. The current president is a Sicilian, the energetic CERN experimenter, Professor Antonino Zichichi. Zichichi is also a passionate popularizer: he runs a science page in a Rome newspaper, is regularly on television, and in the last three years has given an extraordinary 150 hour-long lecture — sometimes four times in one day — in theatres and lecture halls across the country. ("People like contact, they like to ask questions," he says.)

The broad Italian interest in particle physics must be partly due to Zichichi's efforts. For interest there certainly is. At his last lecture in Genoa, there were thousands of people, lucky ones crammed in the largest theatre in the city and 2,000 on the streets outside. All his performances are oversubscribed, he says. His talks are extempore, and in part relate science to religion. "But I am not a religious speaker, I'm a scientist: I speak about science, not religion . . . For just three minutes in an hour I remark that in my discussion of science there is nothing in conflict with religious belief."

In his talks, Zichichi discusses abstruse physical concepts by using familiar analogies: for example he talks of "Abelian and non-Abelian" (commuting and non-commuting) gauge theories by describing ordinary actions where the order of action does not matter and where it does. He talks of space, time, curvature — and the audience appear to lap it up. He would appear to be creating the most extraordinary thing — a popular political constituency for high-energy physics.

Anyhow, that aside, INFN does have its problems. Like all the major state scientific institutions, it has been frozen since 1975 as a *parastato* body, unable to promote on merit or inject talent at high levels. (However at INFN, some of the *parastato* provisions have been ameliorated: staff going abroad, for example, are treated like diplomats, with generous allowances, and

it is possible to take a renewable 5-years' leave of absence to pursue major experiments at foreign laboratories.)

INFN's major laboratory is at Frascati, south of Rome, where the electron synchrotron ADONE gave it a world class position in the early 1970s (just missing the 'hidden charm' J/psi particle because its energy was a hairsbreadth too low). But Frascati suffered badly in 1974 in a split of the laboratory between CNEN (the nuclear power agency, now ENEA) and INFN.

In the division, most of the researchers stayed with INFN, but most of the technicians went to CNEN, guessing that there was more future for them in nuclear power than in high-energy physics. This left INFN's Frascati very much out of balance, says Frascati's machine physics director, Dr Sergio Tazzari, with too many chiefs with too many projects and too few Indians to execute them. "Frascati has 200 or so people — about the size of a single group at the proton-antiproton collider at CERN" says Tazzari. But Frascati is not a single group, but 20, performing experiments at Fermilab and SLAC in the United States, at CERN and in the proton decay facility under Mt Blanc, and at Frascati itself.

Also, under *parastato*, professional salaries at Frascati have fallen disastrously behind industrial ones. The people who built ADONE, says Tazzari, were bright engineers attracted from industry not just because of interest in the work, but because of higher pay. Now the flow is exactly the reverse, with a factor of two salary differential in favour of industry.

Meanwhile, though, the INFN council has developed an ambitious five-year plan, which includes escape from *parastato*. INFN is already building a giant under-

ground laboratory — protected from cosmic rays — under the Apennines near L'Aquila, east of Rome. Dubbed 'Gran Sasso' it will consist of three parallel tunnels 100 metres long and 100 square metres in cross section, to be fully equipped for a variety of experiments like the search for proton decay, solar neutrinos, neutrinos from stellar collapse and so on: "A very complicated structure, very modern, and extremely comfortable" says Zichichi. The five-year plan requests 40,000 million lira (£19 million) to finish the civil engineering of this laboratory.

Then the plan envisages Italy building all the superconducting parts — a major part — of the West German proton-electron collider HERA, to be built at Hamburg. And even more futuristically, INFN is proposing the study of a machine it calls the Eloisatron: European long-range intersecting storage accelerator. It would be 50 km in diameter, slam 25 TeV protons into 25 TeV protons or antiprotons, and rest in the plain of Nardo, near Bari in southern Italy. (This plain, says Zichichi, is the most stable geological area in Europe.) It would be built entirely with superconducting magnets and accelerating cavities, and "should start in five years". At present the costs are prohibitive but they might be brought down to "a couple of gigadollars" (about four times CERN's present LEP project). The accelerator would be European "But Italy is pushing this through — it is to show the European high energy community that we need a big project . . .". The Eloisatron is in competition with the similar American 'desertron' idea, Zichichi says.

The INFN five-year plan requests 1,000 million lira a year to study the Eloisatron. But everything is likely to be delayed by the Italian elections, just called: "I would have said everything would have gone through by the end of this year" says Zichichi. But now he's not so sure. □



An "extremely comfortable" laboratory is being built under the Apennines, safe from cosmic rays

Nuclear power agency

A real power in the land

SOME institutions in Italy work well. One of them, hopes Umberto Colombo, will be the nuclear and alternative energy agency, ENEA. (ENEA is roughly equivalent to, say, the Atomic Energy Authority in Britain, but is without a weapons role.) Colombo, president of ENEA, last year pushed through a reform of CNEN (as ENEA was called until then) which, he hopes, will give it the ability it needs to mastermind Italy's nuclear development — and if he had his way alternative energies and other industry as well.

Colombo works long hours. The highest earning man in the organization, after the director-general, was a man on plenty of overtime: Colombo's driver. Though Colombo tells that story to make a very different point: that the director-general and all the other professional staff of CNEN had ridiculously low rates of pay, way below the rate for comparable jobs in industry. The problem was the old chestnut strangling all the government-controlled scientific institutions in Italy since 1975, *parastato* status (see page 112).

"While CNEN was under *parastato*", says Colombo "many people did two jobs; many people were frustrated; several left; and some were heroes". In reforming CNEN to make ENEA (which has escaped from the *parastato* umbrella), one of Colombo's principles was to raise salaries to no less than 10–15 per cent less than industry. A small differential from industry was reasonable, he argued, because work at CNEN (ENEA) offers more security and more freedom than industry.

So on 29 April this year (parliament passed the reform nearly a year ago) the first 73

ENEA 'managers' were appointed according to Colombo's new rates. By the end of 1984 there will be 180 such appointments. Of this 'commando unit' of Colombo's chosen few, 10–15 per cent will progress on a 'scientific' ladder, and 85–90 per cent on a 'managerial' ladder, "though even that will have a strong scientific content". So ENEA "will reward knowledge, professional ability, scientific and technical standards: that is unique in Italian public institutions", says Colombo.

"We can also form companies" (which was impossible under *parastato*). The companies must be for the development of a technology, and one has already been formed, to deal with nuclear waste. Also, ENEA is permitted to take a minority interest in other companies such as nuclear construction corporations "to check on them from the inside". ENEA, for example, has set up a research and monitoring station at Fiat, to look at the production of pumps for nuclear power stations.

Further, ENEA "can hire experienced people now. Under *parastato* status we could only hire beginners, just two years after their university degrees. Now we can inject knowledge from the top". For the first time in eight years!

ENEA has become "a research and development agency that is in touch with industry, with the objective of improving industry". ENEA can also assign funds to industry, "provided we have control of the quality of work done, and that we follow through with our own research", intermingling with industry. "It's a nice scheme." The limits are that the work has to be related to energy, but since energy contributes a high proportion of the costs of production of many items, this gives Colombo and his agency a lot of scope. "We are really involved with technical change," says Colombo; and it's impossible not to feel that the politicians and industry would be well-advised to let him get more involved.

However, ENEA has a problem with cash. The agency has 18 months to run to the end of its (or rather CNEN's) current five-year plan, on a voted budget of 2.89×10^{12} lira (£1,380 million) "but the government hasn't the money to pay what was allocated". So everyone is subjected to a squeeze. But Colombo had calculated that ENEA needed 10 per cent *more* than the voted budget to complete its five-year programme. "The chances are that we will get a decrease, which will endanger our two major projects." These two projects are CIRENE ('chiraina'), a heavy-water moderated, light-water cooled experimental reactor under construction near Rome, and PEC, a 120 MW (thermal) fast reactor for testing full-scale fuel elements under simulated accident conditions. PEC



Colombo — pushing the Sun and the atom

is being built near Bologna.

Meanwhile, Colombo is preparing the next five-year plan — ENEA's first. It has five main elements. First, says Colombo, "we want ENEA to be even more active", ensuring that the reactors under commercial construction and design, which include Westinghouse derived pressurized water reactors (PWRs), have the best technology and the lowest price, and that Italian industry is making an effective contribution. This implies that ENEA must undertake safety research, quality assurance research, and coordinate its research with industry.

The second part of the new five-year plan will be to put "a heavier accent" on the nuclear fuel cycle. The problems downstream of the reactors "have been almost neglected at the industrial level in Italy". So there will be studies of reprocessing and waste. This is a problem which needs a solution in Italy by the 1990s.

The third element will be fusion: an upgrade of the tokamak at Frascati near Rome.

The fourth target is the completion of CIRENE by early 1985 — provided ENEA is not cash-limited. But why is Italy interested in this heavy water technology, when it has chosen to build PWRs as its main power plants? Because Italy has a deal with the Canadian atomic energy commission to cooperate in marketing CANDU. Colombo feels there is a market for it in the Third World, particularly for the mixed technology of CIRENE. So the question turns round: why, then, did Italy choose the PWR? "We did not want to be isolated from the rest of Europe" says Colombo. "France, Germany, Spain, Belgium and now probably the United Kingdom have gone for it." There is knowledge and operating experience there that it would be stupid to ignore.

"Anyhow, when I came to CNEN in 1979 (from the chemicals giant, Montedison, where he was research director) CNEN was not too severe in assessing the value of the projects. . .". The Italian

Budgets were booming

CNEN (ENEA) finance has boomed in the past five years — particularly in the renewable energies sector. In 1979, the year after Umberto Colombo arrived as president, CNEN spent 1,570 million lira (about £750,000) on renewables. One year later, it was spending ten times the amount (15,616 million lira); in 1981, 23,440 million lira, and last year 46,510 million (£22 million).

In the nuclear sector, CNEN's budget also ballooned, though not quite so much: from 121,655 million lira (£60 million) in 1978, to 259,490 million lira in 1979, 268,550 million lira in 1980, 448,630 million lira in 1981, and 475,550 million lira (£225 million) in 1982 — a budget which saw its greatest rise in 1979 and has remained roughly in touch with inflation since.

However, the future may not be so bright. The CNEN (ENEA) 1980–84 five-year plan was costed at 2,890,000 million lira (about £1,400 million), but a huge deficit in the government budget means that ENEA may not see this plan realized.

nuclear industry would have collapsed entirely if CIRENE and PEC had been abandoned.

"But there is 'value added' — CIRENE and PEC are useful, the latter particularly when the French Rhapsodie, the only comparable research facility, is closing: "by 1987 PEC will be the only reactor in Europe capable of fast reactor fuel testing under severe transients". PEC is to be completed at the end of 1987.

The sixth element of ENEA's forthcoming five-year plan will be "to push energy conservation and renewable energy sources, which are taking off". In energy conservation ENEA is working mainly with industry on projects and on training, on housing, and on coupling energy conservation with renewables, for example using biogas from chicken, pig and cattle farms in a complete rational energy (and waste) plan for the farm. There are already nearly 100 industrial biogas plants in Italy, and ENEA is participating and monitoring the projects to identify the best technologies.

But more than anything in renewables, ENEA is pushing hard on photovoltaic devices, particularly for the sunny south of Italy. The technology has a bigger future than mirror arrays aimed at boilers (like the

European Commission plant in Catania), Colombo believes — these are better for some kinds of process heat. So ENEA is building a laboratory near Naples to study the basic technology of photovoltaics; supporting the two (state-owned) firms in Italy making the devices; and "generating demand" by constructing DELPHOS, a 1 MW solar electric plant in the region of Puglia, in collaboration with the state electric utility, ENEL. The plant will be a testing ground, modular, and using crystalline and amorphous cells, as well as giving power to the grid.

In fact the Naples laboratory is why Colombo left Montedison in 1978, he says. "To me the problem of the Mezzogiorno is best tackled by doing the right research there, rather than building huge plants there that later on turn out to be cathedrals in the desert" — such cathedrals as the petrochemicals plants there, once symbols of hope for the south but now idle and withering in the sunshine. Colombo wanted Montedison, instead, to build a Naples laboratory; he is now constructing one with ENEA. "We had a plan, and an agreement with unions." The Montedison board, however, changed its mind, refused, and Colombo resigned. It was Montedison's loss and ENEA's gain. □

The custom of *lottizzazione*

Party games spoil science

UMERTO Colombo, present president of ENEA (the nuclear energy agency, see opposite) had a bit of bother last year, when he left ENEA — he thought for good — after he had sorted out the agency's problems. He became president of ENI, the state hydrocarbons concern. Now he is back at ENEA, after a row that was front-page news in Italy for five weeks. Modestly, he would rather not talk about the ENI episode (modestly, as he was one of the few people to come out of the episode with his integrity intact.) Now, as then, he wishes ENI well. But the story was that he wanted someone kicked off the board who had been implicated in the Banco Ambrosiano affair. However, this man was supported by the Socialist Party, to whom Colombo had always shown sympathy (although he is not a member). Political accounting (and some danger of blackmail to certain parties) implied that the man had to be appointed to the board. Colombo still said no, and then offered him the Christian Democrats support, so compromising all his political boats. Colombo then left the problem to the prime minister, who reached a political compromise: he moved Colombo back to ENEA, where the post was still vacant, and failed to appoint to the ENI board the man Colombo objected to. The socialists then put in a favoured man as ENI boss.

The political net spreads wide. The Mario Negri pharmacology institute (see page 124) has avoided the problem of political

clientilism so far, but fears it greatly, as government money is taking an ever-greater fraction of its budget. Here is what the chief administrator of the Mario Negri, a cheerful — and principled — man called Professor Alfredo Leonardi, said about the matter: "Italian public funds raise the possibility — I say possibility — of political interference. . . . For instance, if the minister of health belongs to party X, he can say 'you may have some money, but you must recruit some people from this list we'll give you'. The list, of course, includes only scientists from party X. . . . This is done at La Scala!" says Leonardi, raising his voice and smiling broadly "at the Scala, sure! At the RAI [Italian radio and TV] when they want to recruit one professional journalist, they have to recruit three Christian Democrats, two Socialists, one Social Democrat, and one Communist. . . . This is a joke, but it's a tragicomic joke. . . . They have eight people paid for one job: only the professional who is not attached to any political party is doing a good job, the others sit there and get their salary."

"This is the phenomenon of *lottizzazione*. The board of Alfa Romeo (a state-owned company) for example is all *lottizzazione*. The cake is cut according to certain rules: 35 per cent of the Christian Democrats, 10–15 per cent to the Socialists, 10 per cent to the Social Democrats, and, just to keep them quiet, 10 per cent to the Com-

munists. . . (it's better to give a dog a big bone. . .). This is unwritten, but practised. . . . If you don't belong to a party you are cut out; you would never become a member of the board of Alfa Romeo, of all the IRI (the network of state-owned companies). . . ."

So was this Mr Colombo's mistake? "Yes! Mr Colombo, first of all, had the weakness of being connected to the smallest Italian party [the Socialists], and then he went against the will of Craxi [the party leader] — this is the case of Mr Colombo. The members of the board of La Scala are all representative of the political parties, to the point that when they enroll the musicians they ask what political party they belong to — yes, the players! — it's an *opera tragica* and also an *opera buffa*."

But could it affect science? "Yes, in so far as the moment we at Mario Negri had a board established by politicians, they would divide the cake politically. . . . They have tried already to make us change our statutes to put on three members of the government." (The institute successfully resisted, by arguing that three persons already on the board — the president of the University of Milan, the dean of the medical faculty, and the director of the institute of pharmacology — were state employees and therefore representatives of the government.)

"Political influence in Italy is tremendous," says Leonardi "and it's really like a cancer: this *lottizzazione* is one of the worst evils in Italy after fascism. Each political party has to have its share of influence."

Is CNR (the national research council) affected? "Partially, but not so much. Among all the state institutions, it is the most resistant, as the committees are elected by the scientists; and scientists are very independent. . . . we are loose dogs, and the politicians hate loose dogs."

What about the universities, are they politicized? "They are becoming politicized" says Leonardi. The recent recruitment of professors (a massive recruitment since 1980 — see the article on the universities) has been the most politically polluted recruitment ever, since Fascism."

The origin of this great political power is the *stability* of Italian politics, Leonardi argues. Despite the rapid changeover of governments, many of the ministers remain the same, even for a decade and more; in elections, voting patterns in Italy have been almost constant since the war (Christian Democrats 35–38 per cent, Communists 32–35 per cent — but always excluded from government by a coalition of other parties — Socialists 10–12 per cent, Social Democrats 10–12 per cent). The prime minister has only once not been a Christian Democrat (in the last government, Spadolini was a Socialist). Changes of government, although rapid, are very minor in their effects. So Italian political power is actually the most stable of all power structures in Europe. Hence its tentacles have had time to creep everywhere. □

Ministerial office of research and technology

Spearhead of the revolution?

RICERCHE e innovazione — per uscire dalla crisi!, "Research and innovation — to escape from the crisis!" proclaims the title of an Italian journal published last year. "*Recherche — pour sortir la crise!*", "Research — to escape from the crisis!" ran a phrase of the heady, early days of the Mitterrand presidency in France. The author of the French phrase was the then minister of research, Jean-Pierre Chevènement. The author of the Italian book was the then minister of research and technology, socialist Giancarlo Tesini.

But it's a very long political way between Paris and Rome. "We can't possibly apply those Cartesian solutions here", says a seasoned Roman commentator close to Prime Minister Amintore Fanfani. But Fanfani does take science seriously: he established the post of minister of research himself, in the 1960s. And there is now a fairly widespread will in Italy to improve the country's scientific performance, at least as far as it affects technology. And although Italian research ministers don't last for long (12 months is typical), allowing little room for strategic policy-making, this consistent will is enough to get a few things moving.

Moreover, Giancarlo *uscire la crisi* Tesini was preceded as science minister by a man with similar views, social democrat Professor Pier Luigi Romita, a respected Milan engineer. And Tesini was also succeeded by the same man: with the fall and reconstitution of the government last year, Romita returned to the post. Romita has in fact been science minister three times: first in 1972-73, again under Fanfani. So there is a thin thread of continuity in official science policy-making in Italy.

Romita and Tesini would agree, for example, that Italy has a long way to go in improving scientific infrastructures — and in raising the amount of research that's done. Total research and development spending in Italy may have just scraped above one per cent of the gross national product (GNP) — less than half the typical figure of a Western industrialized country, and low even allowing for the fact that Italy does very little defence research. And the management of science, and its harnessing to industry and the economy, is loose, patchy, and sometimes even anarchic.

"Italian development has always been a little behind other countries", Romita says by way of explanation. "Our national unity came late (1871), and also our industrial revolution" (which was based partly on hydropower in the northern Alpine valleys). So the understanding in Italy that research was connected to economic development came late too.

"Also, our industrialists and entrepreneurs have been short-sighted — they sought immediate benefits rather than

establishing the conditions of a continuous development". Even so, says Romita, Italy ranks seventh in the world league of industrialized powers — entirely because of the great efforts made by government and industry in the 1960s: the Italian miracle, it was dubbed at the time.

"But this was typical of an emerging country", says Romita, "we devoted resources to industry which in a developed country would be devoted to other things, like social welfare." Labour costs then were also very low compared with those of other Western countries. So Italy developed very rapidly in the 1960s, and production reached the level of an industrialized country. "But then new problems arise." The self-consciousness of the working class, the strength of the trade unions, welfare payments, wages — all



Outgoing Prime Minister Fanfani established the post of research minister

these increased rapidly. "So now our economy is the economy of an industrialized state, with all the advantages and disadvantages that entails."

"It's exactly at this point that the necessity of research and development becomes clear, and becomes the basis of further development. This consciousness has been growing in Italy since the end of the 1960s — consciousness that research and innovation has to play a much more important part than before."

At the end of the 1970s, Italy's private and public financing of research and development reached 0.9-1.0 per cent of GNP; and at the end of 1982, 1.2 per cent. "We have crossed the magic barrier of 1 per cent, and we are continuing to invest in spite of — and perhaps because of — the crisis."

Italy has no natural resources; its

balance of payments is in the red, particularly in oil but also in food; it must be a transformation economy, adding value; so the country must do more research and innovation, "as must all industrialized countries that want to remain industrialized countries".

Ex-minister Tesini had been pursuing a "law for science", defining budgets and a programmed development like France; Romita differs only in wanting to approach the same ends in a more piecemeal fashion. "A general bill for science would not be easy: there are so many different bodies involved, different authorities, different laws." What the science minister can do is to establish a clear view of the objectives; of the resources required; and then aim to achieve them "by a number of legal initiatives".

For example, the budget of the Consiglio Nazionale delle Ricerche (CNR) — the principal government-controlled research body — could be programmed better (in the last two years it rose 30 per cent, and then fell 20 per cent); there could be programmed budgets for the Istituto Nazionale di Fisica Nucleare (INFN) — for high energy and basic nuclear physics — and the two-year-old Consiglio Nazionale di Università which distributes research money to universities; and for government financing of industrial research. "Annually or biennially, we should establish a clear plan of research and innovation, determine the total amount we are going to spend, and then distribute that through the various laws (for the different research bodies and government research initiatives)".

But in Italy, even such a slight degree of planning is a tall order. "We don't have such a structure now," says Romita. "Research has everything to offer Italy," but Italy cannot cope with it yet. "The main problem we have to face to overcome our crisis is the way we organize our society, our way of living and governing."

Ex-minister Tesini agrees, particularly regarding CNR, which is like the various research councils of another country all rolled into one: "the situation at the CNR is very confused", he says. It combines policy making and administration, and has a very meagre administration and little outside control, although it should be the "zip" between research and industry. The reorganization of CNR was the principal problem of the minister of research, Tesini felt.

Romita, on the other hand, puts the emphasis on reorganizing and re-staffing the 'science ministry'. Because in fact, it is not a ministry at all, but merely an office of the prime minister: the minister is 'minister for the coordination of scientific research and technology', and really he has few direct powers. He is a minister "without portfolio". And in fact, according to one cynical civil servant, Italian science ministers generally struggle so hard during their short period of office to establish a portfolio, they have little time to do

anything practical.

Romita is no exception in his interest in a portfolio. But he can hardly be blamed. Almost wringing his hands, he says the key is not so much to have direct control of the science budget, as to have the bare minimum of a permanent staff. Romita has 200 people working for him in his block of offices overlooking the Tiber, but they are all temporary, on loan from other departments. They are often uninterested in the job, and unqualified.

But, Romita claims, the science minister's staff have unique tasks. They must:

- Evaluate the broad needs of the nation in research in different fields, such as energy, electronics, telematics and bioengineering.
- Ascertain and evaluate what research is being done.
- Evaluate the results of completed research.

- In the light of the above, plan a detailed research programme and distribute it among the various research organizations.

Such functions are "completely different from any other work" in government, says Romita; but at every change of minister, the whole staff may change. Fighting against this, Tesini and Romita have tried to keep "the minimum" of continuity between their successive ministries. "luckily", the ministry maintains many advisers — often university professors — who act unpaid on committees; but this is no substitute for a permanent staff.

Also, in applied research, the ministerial task is becoming more and more one of handling a substantial budget, and a permanent staff is necessary to ensure it is properly managed. Last year, Tesini steered through a bill which had been set in motion by Romita before him, which in the biennium 1982–83 gave the minister 1,700,000 million lira (£810 million) to spend on applied research: the minister's first real budget.

The money is geared, in part, to 'national research plans' in strategic areas, and is spent by means of research contracts with research operators (such as CNR or the universities), somewhat on the lines of the Rothschild principle in the United Kingdom. But instead of a distribution of customers in different ministries, this new Italian system is more centralized: Romita spends the money.

Or rather, he makes propositions within the agreed limits to the interministerial economic planning committee, the all-powerful CIPE (pronounced cheepay). This committee, and the related committee for industry, CIPI, have had the habit of establishing national plans for different sectors of the economy (which, incidentally, is around 50 per cent nationalized); and to these plans, according to Romita's initiative, are now associated corresponding 'national research plans'. A current example is energy research, which is planned through a ministerial advisory committee



An Italian government. They change so quickly the faces blur . . .

in relation to the national energy plan. The national space plan and its associated research are older, outside this scheme and its budget. For the future, CIPE has attached high priority to electronics, aviation, space projects not covered by the national plan, biotechnology and chemistry.

Outside the national research plans, and beyond the space plan, Romita is spending on applied research projects of CNR (so-called *progetti finalizzati*, short-term research geared to a defined practical end), and research requested by small and medium industry (where Romita contracts out and pays for research that the small firms cannot do themselves). The spending was planned to be 700,000 million lira in 1982, and 1,000,000 million lira in 1983; the first target was not reached (money never comes when it's wanted in Italy, even if the sum is agreed) and there are now objections to the second sum, which could end up being halved; but nevertheless the amount is new money and not negligible.

"This is the first time my office can allot funds", says Romita. "If we now have a permanent personnel structure, we will have a real ministry." It's logical, but there's a drawback: the man in the street is rightly suspicious of a generally inefficient and slow Italian bureaucracy, and there is public pressure to reduce the number of ministries rather than increase it. Nevertheless, in the industrial sector there is interest in creating the ministry, and there is also a feeling that a permanent ministry is necessary to handle international relations in science.

(Amusingly, on a recent short-notice visit to Britain, Romita was caught out because he could find nobody in Rome who had knowledge of British science policy and management. The post of science attache in London has been unfilled, to date, for nearly a year, largely because of confusion over determining the salary for the

job; so there was nobody to advise in London either. The previous London attache was occupied with a conference in Pisa, and Romita was reduced to copying out, on an A4 copier, sections of *Nature's* 2-foot by 3-foot chart on the organization of science in Britain. These were duly pinned together and formed Romita's main source of information on the British scene.)

But Romita's emphasis on establishing a ministerial structure does not mean he has forgotten CNR. Between his posts as minister, he was chairman of the parliamentary deputies' committee for education and research, which oversees most laws affecting science (there is a roughly parallel committee in the Senate), and there one of his tasks was to prepare a reform of CNR. "I am determined to carry on with this reform", he says.

However, in his committee, Romita was indirectly responsible for delaying it, something he may now regret. For a previous minister had proposed both the strengthening of the science "ministry" and a CNR reform; but the reform was to be "by decree", which is to say that the power of detailed reform, within certain guidelines, would have been delegated by the parliament to government. Romita, in his parliamentary guise, fought for parliamentary definition of the whole law. There the matter rests, with no agreed text, although there are now outline proposals from different parties and the government. (See below. The most detailed proposal comes, as it happens, from the Communist Party.)

And what were Romita's main tasks for the year, before the fall of the government two weeks ago? To complete the preparation of the national research plans, reform CNR, and establish a ministerial structure for science, he says. But Italy being Italy, he now seems unlikely, with the best will in the world, to complete all of those tasks. □

Mario Negri pharmacology institute

American inspiration in Milan

THE Mario Negri Institute for Research in Pharmacology in Milan is the tops. It is consistently in the US National Institutes of Health (NIH) list of top ten most-funded foreign institutes (at least twice it has taken top rank) and its director, Professor Silvio Garattini, is one of the only four Italian scientists to have reached the Institute of Scientific Information's list of 1,000 most-cited scientists (see page 113).

And yet its story is almost a fairytale, a near-impossible constellation of the right people at the right moment. The "fairy godmother" took the shape of Mr Mario Negri, successful jeweller who was very interested in medicine and — late in life — became a pharmaceutical entrepreneur. He wanted to put his profits into research.

"We're a little bit atypical in Italy," says Garattini, guilty of British understatement. (The institute is probably unique, but it shows a way in which research really can be done in Italy.) The idea began 26 years ago, when Garattini, then at the department of pharmacology in the University of Milan, paid a visit to the United States. "I quite quickly got the idea that we were amateurs," he says. He told his 25-strong group so and suggested either they all went to the United States, or tried to do something new — modern, efficient, large enough — in Italy. They chose the latter. Garattini had also been impressed by the many kinds of support for institutions in the United States. He met Mario Negri the year he went to America (1957), when Negri came to the university for advice on research. Says Professor Alfredo Leonardi, then a member of his group and now general secretary of the Mario Negri institute: "Negri was very enthused with us, and dissatisfied with the other university groups. . . . We talked a lot with him, showing him drug tissue cultures, and the effect of his drugs on cancer cells."

The university was much too bureaucratic to get anything done, says Leonardi (the building they worked in dated from 1800), so they went to Negri with a proposal for an independent institute. Garattini and Leonardi were then just 30 years old: too young, said Negri. "Then we had a couple of difficult years." They brushed up their English (they were the only two to speak it in the group). "Other Italian professors did not speak English at that time," says Leonardi. "When I became an associate professor in 1960, all my publications in English were ignored."

Then, in 1960, Mario Negri developed cancer of the liver. A week before he died, Garattini went to him: Negri said "make sure everything goes as we discussed". Garattini was certain that meant that Negri had put them in his will, and so it happened — they were left 1,000 million lira — 100

million in cash and 900 million in the form of stocks in the Negri pharmaceutical company. The will was very detailed, defining the rules of the new institute, that it would do research and education, and run conferences just as the young scientists had specified in a long memorandum two years before, and naming Silvio Garattini explicitly as director. There were legal wrangles, though, as the university tried to absorb the institute, and it would have done if Garattini had not been named as director. So it was a long time after Negri died that Garattini was called to the solicitors. Leonardi stayed in the street below. "I knew it was all right when he waved a piece of paper from the window."

"The affair shook the university," says Garattini "and four years later they built a beautiful laboratory for 250 people". So with the Mario Negri institute itself, the result has been to make Milan a very important centre of pharmacology in Europe, both in the research and production of drugs.

Then Garattini's group had another decision to make: to invest the money and use only the interest — which would have paid the costs of the whole group — or spend it quickly to set up a big institute, in the hope that research grants and maybe other endowments would later keep the place going. They chose the riskier course; got a mortgage from the (independent) regional development bank CARIPLO, money for a library from the Pfeiffer Foundation, £70,000 from the Wellcome Trust for all the basic large experimental equipment (Negri had been Wellcome's agent in Italy), and by the end of 1961 when the building began the group already had NIH research contracts to begin in February 1963. "We were crazy" says Leonardi. "I would do it again, but be less crazy."

The institute began in 1963 with 16 scientists. In 1982 it had 154. In the first year, it produced 4 papers, one book, and one set of conference proceedings. In 1981, it produced 153 papers, 26 general publications (including popular writing), 8 books and 3 proceedings of conferences. In total, over 18 years up to 1981, the Mario

Negri has been responsible for 1,488 papers, 163 general publications, 51 books, and 41 conference proceedings. Some 7,000 foreign scientists have attended its conferences, and there have been over 200 foreign scientists working at the institute, with an average stay of 10 months. Another 500 young people have been awarded a special 'degree' for training in research at the institute (a kind of PhD when Italian universities had no such thing), and in 20 years 1,000 young people have passed through, to almost certain productive employment.

These figures could go on and on, but there's another interesting aspect of the Mario Negri: the degree of ethical thinking behind the institute. The research covers three main areas: cancer chemotherapy, psychotropic drugs, and cardiovascular drugs. But the interest of the Mario Negri is not in developing new drugs: "even if we do, we don't take out patents — we publish everything". Rather, it is to understand the mechanism of drug action, both in therapeutic and toxic effects. Research is done in collaboration with industry "but we need to have a common interest. We are not a contract research institute like Battelle: our object is basic research in pharmacology," says Garattini. Also, "we go from the molecular level to the human being". Drugs are followed through in their use in hospitals and by general practitioners. "When a drug goes onto the market, that's only the beginning," says Garattini "of assessing the balance of risk and benefit that ultimately determines the value of the drug". The institute has several collaborations with hospitals in Lombardy (the region of Milan), conducts clinical trials, and has a large epidemiological group.

For Garattini, the ultimate value of a drug can be measured by the change it causes in pathology. Polio vaccine was a good product because it reduced the number of polio cases. "For many other drugs there is no correspondence: you have the drugs but the pathology does not change." "We have too many drugs on the market: too many serve no purpose." Garattini adds "This is the value of being an independent institute — you can say what must be said".

"Too frequently medicine has been a way for corporations to get their share, rather than be at the service of the sick. . . . and we always hear of the rights of doctors,



Negri director Garattini (left) and general secretary Leonardi

nurses, industry, pharmacists, bureaucrats, administrators, and the needs of the economy and so on; from time to time it is very important to remind ourselves and others that this is the purpose of medicine."

So the institute is involved in the popularization of its ideas, in newspapers and broadcasting, "telling the public what science can and cannot do". It has also developed a formulary (a list of drugs with descriptions of uses, doses and effects) which has "relatively few drugs" compared with the large numbers on the market. The list is now in use in the hospitals of Lombardy. Also, the Negri has now gone to the second edition of a formulary for general practitioners "where we tell them about the drugs that are active, and of course there are very few in relation to the available drugs". Garattini and his clinical pharmacology colleague, Dr Giovanni Tognoni, were also in the World Health Organization panel that elaborated the concept of 'essential drugs' for the Third World.

"We are also interested in improving the use of drugs," says Garattini, so the institute (and particularly Tognoni's group) works with medical practitioners on these problems. For example, they have developed a protocol with 700 general practitioners to investigate the best ways of controlling hypertension in old people, which ultimately will lead to data on several thousands of patients. Also, Tognoni has organized 35 neurosurgeons to collect trauma data "which could become the largest data set in the world" on the problem of head injuries.

To organize such a thing, and also clinical trials, is not a trivial exercise, Tognoni says, as the idea of individual research has been dominant in Italy, as in Germany: collective research (where the results and the glory are shared) is the exception rather than the rule". The Mario Negri has advanced this idea of shared results.

The next task of the institute is educational, to help young people become scientists. The Mario Negri three-year research degree is awarded by the regional authority, after an examination before an external committee when the candidate must discuss the scientific papers he or she has published and give a 40-minute lecture on their work. It is effectively a PhD in pharmacology, but it goes by the rather slight title of 'qualified in pharmacology'. "The universities are just starting their 'PhDs', but we've had them for years," says Garattini.

The institute is still expanding, adding 1,500 square metres of laboratory space on site in Milan, and building a research centre in Bergamo, 25 km away, to be ready before the end of this year, to give a closer interaction with clinicians in hospitals — with whom they already conduct much research. But having the laboratory in Bergamo will mean two-way communi-

cation: clinicians will be able to work out their own ideas in the laboratory as well as researchers working with patients. The laboratory will house 100 people.

Then there is a major exercise in the South of Italy, in Abruzzo, with money from the Cassa di Mezzogiorno, the government fund for development in the South: Negri South, they call it. "Intelligence is randomly distributed," says Garattini "so you can find it all over the world". The north of Italy simply has had more opportunities. "When we had to select 16 researchers for Negri South, we had 200 applications of which 60 were very good — so good that selection became essentially arbitrary."

Money at the Negri, however, is getting tight as the recession bites on Italy — and as President Reagan savages the basic science budget in the United States, affecting the NIH contribution to the institute drastically. "Of course, American science must have the priority," says Leonardi "but the reduction has been very great". At the end of the 1970s, the institute was supported by

roughly two-thirds Italian funding, and one third foreign. In 1981 foreign funding collapsed to a fifth, largely because of a near-withdrawal by NIH, whose contribution fell from \$1 million to \$0.2 million.

"We much prefer American funds," says Leonardi "because they pay 'with decency', which means to say on time. Italian public funds always arrive very late, which is terrible in time of inflation. Also there can be political interference" (see the item on *lottizzazione*, page 121).

The total funding of the institute has continued to rise, however, because of an increase in grants from the Italian national research council (CNR) and in private donations. These have come especially (500 million lira) from the cancer charity, the Italian association for cancer research (AIRC) which has, says Leonardi, undergone a transformation since Mr Venosta, formerly head of Pirelli, the tyre firm, took control of the association, raising its income from 1,000 million lira to 13,000 million lira in five years. But that's another story. □

International Institute for Genetics and Biophysics

Ambitious despite the problems

NAPLES, they say, is the city where all the problems of Italy are greatest. The city is beloved and decaying: "It's a marvellous place to live in," said one young researcher at the university "but it's impossible to work here". Regretfully, he is now looking for a job further north, or abroad.

Yet perhaps half of Italy's real contributions to molecular genetics come out of Naples. Or at least, so one Neapolitan says. Professor Francesco Blasi runs the largest CNR institute, the International Institute for Genetics and Biophysics (IIGB), whose focus is human molecular genetics. IIGB is based in Naples. There maybe three groups further north, in Rome and Milan, "doing really new things" in molecular genetics, Blasi says, but there are "several groups" in Naples in the forefront of the subject.

And foremost among them is IIGB, which is the remnant of an even bigger dream, the dream of Professor Buzzati Traverso for an Italo-American postgraduate 'university of molecular biology', which was going to attract the best minds in the field. Buzzati Traverso died only two weeks ago; on the next page we outline his vision, and how Italy failed him.

Back in the present, Blasi has plenty of ideas about what ails science in Italy, and particularly CNR. First, though, after three years at the US National Institutes of Health (NIH) and four in West Germany, why is Blasi doing research in Italy at all (apart from the fact that he is Italian)? When he went to NIH, he had decided to leave his country. He returned because of the "astronomical numbers" and tremendous enthusiasm of Italian students for research, he says, enthusiasm greater than

he could find at NIH. Moreover, IIGB does "excellent work" now, says Lucio Luzzatto, ex-director of the institute and now at the Hammersmith Hospital, London. And "there's much more science made in Italy than you would expect, given the situation".

Blasi's diagnosis of "the situation" is summed up in six points:

- The major problem, Blasi says, is that CNR is suffocated in the mass of rules and regulations because it is *parastato* (a 'para-state' organization, see page 112). CNR regulations are identical with those controlling budgets and careers in the national football pools service, leading to anomalies like being unable — officially — to let a visitor use the canteen. "These stupid things make our life complicated."

- The next problem is money: "I calculate that our budget per scientist at IIGB is just 1–2 per cent of the budget per scientist at the European Molecular Biology Laboratory." Admitting, though, that EMBL is an exceptionally well-found laboratory, he insists that IIGB — and other Italian institutes — really need a two- to threefold increase. For example, the IIGB budget is about 1,000 million lira (£480,000) a year including administration, but excluding salaries for its staff of 150. The money must pay for all the basics "including the dustmen". The resulting cash for equipment and materials per scientist works out at "about \$10,000 [£6,000]". Then there is another 300 million lira (£140,000) in grants (all *progetti finalizzati*, see page 114) to certain groups. Only members of these lucky groups are able to get up to a respectable \$30,000–50,000 per

● Finally: the bureaucracy. The political situation has completely changed since the 1960s, says Blasi. Then Italy was a private enterprise country. Now it is state-controlled. While that has its advantages, "the bureaucracy is terrible, it stops everything and it helps only what is worst . . .". Bureaucracy permeates the whole university and scientific structure of Italy, says Blasi. And as for the reform of the CNR, the planned reorganization that might bring it out of the *parastato* limbo: "the law for the universities [voted in 1980] took 30 years! These things can take for ever." □

The first curriculum was supposed to start in 1967. It was postponed to 1968, then 1969: and then the student revolution struck. "The students," Schreil claims "found an alliance with the most fascistic,

© 1983 Nature Publishing Group



Schreil — mourns Buzzati Traverso

most extreme elements of the University of Naples. They teamed up". The slogans of "Help Laos, throw the Americans out of Vietnam, help the Khmer Rouges", were merged with "throw those university barons out of the LIGB!". The extreme right subsidized and helped the extreme left, Schreil says: "Everybody knew it."

But weren't the University of Naples people also 'barons', and so equally likely to suffer the scorn of the students? Yes, says Schreil, but the successful university tactic was to divert the attention of the students to LIGB. Presumably, given the strong American content of the laboratory, and the anti-American feeling of the times, this was not difficult.

So in the spring of 1969, the LIGB was occupied for three months by students and some of its staff, technicians and 10 Italian research fellows who, says Schreil, wanted permanent positions. (Buzzati's idea had been that the fellows would work three years at LIGB, then go to the United States, and then return — if they wanted to, and were good enough).

But Buzzati Traverso and his closer colleagues were not interested in politics, whereas — according to the Euratom observer — the occupiers "were extremely able politically". So "they made rings round the professors". Buzzati asked CNR what to do — should he call the police? — but he received no assistance. He gave CNR an undated letter of resignation which, at a certain point, CNR used. From then the dream was over. Buzzati Traverso was sent immediately to UNESCO as an assistant director-general, where he stayed four years; and he spent the last decade of his life campaigning against nuclear weapons.

"Buzzati made one last attempt to save the institute," says Schreil. Just before Buzzati's 'resignation', and with the secret letter already on the CNR president's desk, Schreil had a dinner party which included the German ambassador, the German consul in Naples, people from Euratom and other influential persons — and Buzzati. "Buzzati had given up," says Schreil. "He said 'I have given everything to these people — they don't know what they want — I might withdraw'. But everyone at the party pressed him very hard to keep trying, and

though he had definitely abandoned Naples, he said he would try Rome. Hurred plans were made to transfer the LIGB to a place in Professor Levi Montalcini's institute in Rome. But in a very short time she came under very heavy pressure — now from the University of Rome — to withdraw from the scheme. For Buzzati this was the last straw. Two weeks after the Naples dinner, he 'resigned'.

Now the remnant IIGB continues, and has even become the largest and a successful, CNR institute (see above). But even the present director, Professor Francesco Blasi, admits its total output declined since the early years, when there was a number of foreigners in the laboratory. "But if you consider the Italian element, that's not true . . .". While the foreign element has all but vanished — Professor Schreil is the sole exception —

the Italian contribution has risen steadily over the 1960s, and also stimulated the university to take molecular biology seriously — so that Naples is now, after all, the principal city for molecular biology in Italy (or so Professor Blasi believes). That, at least in part, must be regarded as an achievement of Buzzati Traverso.

However, today's IIGB is not a 'studium'. Some say Buzzati Traverso was naive to think that he could ever have succeeded with his plan in Italy. His American friends, who knew the country, had warned him repeatedly of the danger of some kind of coup. But Schreil says he was not naive: "He just didn't care for details. He didn't want to get involved in the day-to-day affairs". For him the student revolution and the university opposition were such "day-to-day affairs"; eventually they overtook him. □

The Naples Zoological Station

The Woods Hole of Europe?

"I AM going to establish in Naples a large aquarium for the public" wrote Anton Dohrn, German zoologist and Darwinist, in 1870. "A further floor will be reserved for the scientists . . . the land behind the Villa Reale down by the sea is very cheap and the building stone can be obtained nearby. The tuff for the grottoes can be brought in masses from Vesuvius, the seawater is always fresh in front of the door, and the animals occur by the million in the sea; all can be done very cheaply".

The result was the Stazione Zoologica, the zoological station of Naples, which became a magnet for European (and American) biologists for almost a century. The station has had close links with Britain: Dohrn met T.H. Huxley in 1867, and became a warm friend of family; Huxley introduced Dohrn to Darwin, who was one of the first visitors to the station after research began in 1873; and right up to the present J.Z. Young uses the abundant squid and octopus of the bay of Naples to perform his experiments on nervous conduction at the station.

The history of the station is golden and delightful — but there is hardly room here to give even a thumbnail sketch of it: we are concerned with more recent times. There is interesting reading, though, in somewhat earlier editions of *Nature*: Dohrn also knew Norman Lockyer, the first editor of *Nature*, and he wrote a number of articles for Lockyer between 1871 and 1875. Enough perhaps to say that 18 Nobel laureates have worked there, from Van't Hoff to Karl von Frisch. They were attracted by many things, which grew up together: first, of course, abundant fauna in the bay, brought fresh to the station (which was right on the beach); then fishermen, paid for by the station, who came to know the scientists' every needs, and the fine detail of the life of the bay; the

development of superb staining and cutting methods, and preservation techniques by a technician (named Bianco); the best microscopes (Zeiss, another friend of Dohrn, tried out his prototypes at the station); the eventual atmosphere of a continuous scientific conference; a good library (now superb); the stimulus and entrepreneurship of Dohrn himself, who could charm and befried people from kings to fishermen; and the excitement — in the early days — of applying the new Darwinian theories.

Some of these special reasons for the station have now vanished, however, and the number of visiting scientists has dwindled from a peak of over 200 in 1960, to only a few tens today. But the station is still doing good work, the fauna is still rich despite the patchy pollution of the bay, and the staff of the station and still-devoted visitors are thinking of some kind of revival.

The station stayed in the Dohrn family for three generations — the last, Peter Dohrn, resigned with his governing council in the student revolution of 1968-69, when the students and technicians decried the station as a 'scientific hotel' for foreigners, of no benefit to Italians or Italy.

There is some truth in that: 1968-69 was "a palace revolt" at the station, says American Dr Carmello Thomas — who has been brought in to restart the marine botany division of the station (it had been defunct for five years). The budget of the station was ailing, and staff did not know from month to month whether they would be paid.

But since 1969, and particularly since December 1982, which brought a new law for the station, it has been thoroughly Italian, a National Institute of Italy, putting it in the same situation as the astronomical community (in their 'observatories'), embraced at last by the very *parastatale* condition that is choking the



Naples Zoological Station, from a print dated 1873, the year the station opened

CNR (the national research council). "We had to get on the bus in order to get off" says the station's last director, Professor Alberto Monroy, hopefully.

The advantage of the deal? Money, says Gioacchino De Vivo, administrator of the station. After the departure of Peter Dohrn, the station was funded *ad hoc* by occasional acts of parliament, a very precarious state of affairs. Up to the end of last year the station's fixed income was 1,700 million lira (£180,000), a sum determined in parliament in 1977. That was a small sum for the station even then, and Italian inflation over five years meant that the real needs of the station in 1983 were 4,000 million lira. The new act of parliament gives the station this sum (in principle), with, at last, a regular means of determining the next budget.

But what to do? Vice-President of the station now is University of Naples professor Luigi Mendia. There are few visitors now, he says, because proximity to the bay is not as important as before (except for people like J.Z. Young, who uses hundreds of octopi in a season): "you can receive materials from Naples in 24 hours by plane in Vienna" says Mendia. Also, there are now many similar laboratories; and the equipment and techniques are now used throughout Europe, Japan and the United States. There will be no appeal in the station, he believes, until it develops strong research programmes of its own which will attract visitors in their own right. Hence the division of the station, now, into departments: ecology, biochemistry, cell biology, neurophysiology, botany, plus the public aquarium, museum and library. The latter, says Mendia, is excellent, "a lighthouse", and it is funded, as a remnant of the German connection, by the Deutsche Forschungsgemeinschaft (DFG) in Germany. DFG also pays the salary of the archivist of the station, Christiane Groeben.

Alberto Monroy, cell biologist, would like to see the station regain an interna-

tional flavour in another way. The Stazione Zoologica used to be the world's marine biological congress, says Monroy. Many Japanese come to Naples, but in general "the image of the station has decayed". The accolade must now go to Woods Hole, Massachusetts. But the Naples Station could still become a "Woods Hole for Europe", with semi-continuous congresses and courses mixed with research. "The great strength of Woods Hole is the courses" says Monroy. "If Naples could run such courses it would really re-acquire its leading position in Europe . . . They should be no less than six weeks long, practical courses in which the students do the experiments, engage in some small project."

Naples has advantages over Woods Hole: the latter institute, on the cold Atlantic, is used mostly only in the summer, but useful species are available the year round at Naples. But there is one great problem with the station, given the present interests and fashions of biology: it is not equipped, and cannot cheaply be equipped, to use modern molecular techniques. "For example, we [station staff] have started some work with monoclonal antibodies. It would be nice to do it here . . . but you need to work under sterile conditions, which is very difficult where there are moulds coming out from the walls, from everywhere!"

So Monroy does his molecular work in Sienna "just five hours away by car. OK it's a long way. But by American standards it's just round the corner, and everything is there". The work cannot be done elsewhere in Naples — reputedly the strongest city in Italy for molecular biology — because of lack of interest, among other Neapolitan geneticists, in using marine animals. Says Professor Blasi, director of the International Institute of Genetics and Biophysics, just a few minutes — not five hours — away: "You do not collaborate on techniques; you collaborate on problems . . . Their objective is fishes and so on . . .

We are working very much on human genes". Despite this lack of overlap, there is some collaboration between institute and station over the sea-urchin, however.

As for the "Woods Hole" courses, there is another problem the station must face. There is little room for expansion on the site, unlike Woods Hole. There is now a wide highway between the station and the sea, and around the station is a public park (a fragment of the Royal Garden in which the station was originally established). But there is an island, the beautiful Isle of Ischia, out in the bay, where the Dohrns built a lovely house: some research (on benthic ecology) is done there, and the station runs some courses, but they are theoretical. So the question arises of equipping the place for experiments, of paying staff: in short, of money. Also, Naples is now very short of housing, even apart from the earthquake: "and you can't think of running a course without decent accommodation". So is there any solution? "I don't know. I'm very concerned about it," says Monroy.

But finally, let a visitor, Professor Berndt Hägström of Stockholm, who has been visiting the station for 30 years, say why he still comes, despite all the problems. Take the famed pollution of the bay — does that not keep him away? The bay is cleaner than 30 years ago, says Hägström. The natural conditions are as good as 100 years ago: the ecologists say 98–99 per cent of the species described in Anton Dohrn's time are still present. "I've given up many things to come here because of the excellent material." In fact, conversely, Hägström comes because of the extreme sensitivity of certain species (such as sea urchins) to pollutants, carcinogens, mutagens and so on, in concentrations as little as 10^{-10} M. Thus he has discovered, for example, that the action of thalidomide is to make a single point mutation (but he cannot publish the result, because of his full-time employment by the Swedish government pharmaceutical company Kabi, who are sensitive about such things).

Also, Hägström boasts a particular philosophy of biology, which favours work at the station: "Marine organisms like sea urchins or fishes are natural populations. If you take white mice, or rabbits or rats — they are not natural populations, they are first of all albinos: and in the 1920s they knew very well that naturally there were at least 17 alleles for skin colour — so the albinos are depleted. But nobody talks about it . . . Also ICI did some experiments on albino mice, and found that over nine years the natural background of cancers rose from 10 per cent to over 80 per cent, so they couldn't see any difference among the mice given carcinogens." Similar problems affect cell lines. This is why so many substances pass laboratory tests, Hägström believes. "So I admire the natural populations . . . they are in balance" says Hägström. As a consequence "I'm very much impressed at the real possibilities we still have here around the gulf of Naples". □

Price of defining common time

Defining time in relatively moving frames of reference so as to preserve the commonsense notion of simultaneity is possible, but only at an unacceptable price.

THOSE who from time to time send to *Nature*, as manuscripts intended for publication, proofs that the special theory of relativity must be wrong will in future, at least in the first instance, be referred to the issue of *Il Nuovo Cimento* (74B, 67; 1983) in which Dr J. P. Hsu from the National Aeronautics and Space Administration's Goddard Space Flight Center works out some of the consequences of rejecting the notion that the velocity of light is the same in all frames of reference moving relative to each other (or, more strictly, in all inertial frames). For although Dr Hsu's professed objective is to demonstrate that some "fundamental" constants are more fundamental than others, he embarks on his argument by throwing away the assumption that the velocity of light is constant (and isotropic) for the sake of a system for measuring time which satisfies one of the rudimentary goals of the anti-relativists — doing away with problems of simultaneity. In passing, and inevitably, he shows that the game cannot be worth the candle.

Hsu's starting point, explicitly and emphatically not that of an anti-relativist, is that it should be philosophically permissible to require a system for the measurement of what he calls "common time" that should be valid in all relatively moving frames of reference. The procedure is straightforward, and follows Einstein's original discussion of the problem. Within one frame of reference, a system of identical clocks ticking at what is supposed to be the same rate can easily be synchronized by means of light signals exchanged between observers attending the various instruments. Similarly, a population of clocks in a relatively moving frame of reference can be synchronized among themselves, and then there is nothing to prevent those in charge of the second system from adjusting the rate at which each in the set of moving clocks ticks so that the two systems always tell the same time. And so on, to other systems moving with different relative velocity. So there is no practical obstacle to the construction of a system of common time in which people's commonsense expectations will not be offended. The other consequences, however, will be less palatable.

For the purposes of argument, it is sufficient to consider two linear frames of reference S and S' , with the second moving relative to the first with velocity v . In standard special relativity, the transformation from one system to another is easily ac-

complished by the relations $x' = \gamma(x - vt)$ and $ct' = \gamma(ct - \beta x)$ where x, x' and t, t' are respectively the coordinates and times in the two frames, where c is the velocity of light, β is the ratio v/c and γ is the familiar $(1 - v^2/c^2)^{-1/2}$. Implicit in these relations is the assumption that the velocity of light is c in both frames of reference. If, however, it is the time that remains unchanged, the standard equation for the transformation of time must be changed into a form that Hsu writes as $b't = \gamma(ct - \gamma x)$, where b' is neither the velocity of light in the second frame nor even, for that matter, a constant. But it does turn out that the velocity of light in the second frame is given by $c' = \gamma(c - v)$, from which it is in no sense surprising that, on the assumption that the velocity of light is isotropic (the same in both directions) in the first frame, it turns out to be anisotropic in the second.

Logically, this is a serious stumbling block for anti-relativists who may have chosen to follow Hsu thus far, for if it is accepted that either frame can be chosen as a starting point for a system of universally synchronous clocks, the notion that the velocity of light in all other frames should be anisotropic is much more absurd than the usual anti-relativists' complaint about the twin "paradox". As it happens, Hsu does show that the anisotropy of the velocity of light in moving frames of reference should not affect measurements of the velocity of light as obtained from a round-trip measurement as in the Michelson-Morley experiment, whose null result is therefore not an absolute bar to attempts to construct a system for the measurement of time in which the difficulties of simultaneity simply melt away.

Hsu, of course, is bent on frying other fish and, in particular, on finding criteria by which the different fundamental constants can be distinguished from each other. For this purpose, he manipulates the laws of motion for a charged particle in an electromagnetic field in such a way that they remain invariant when transformed according to the rules appropriate to common time, not standard special relativity. Hsu chooses to modify the action integral or, which comes to the same thing, the lagrangian which is its integrand, but the job could no doubt be done in other ways.

One obvious consequence, however, stems from the mere form of the equations of motion of, say, an electron in an electromagnetic field, where the particle is coupled to the field by means of the factor

e/c , where e is the charge of the electron in electrostatic units. So what happens if the velocity of light changes from one frame of reference to another? Plainly it is easiest to formulate equations of motion that remain invariant if the factor by means of which the components of the electromagnetic four-potential are multiplied, e/c , remains invariant, and Hsu shows that this is indeed a consequence of his argument. But e/c is nothing but the charge of the electron in electromagnetic units, $\bar{e}$, which is thus singled out as a kind of preferred fundamental constant. By a similarly straightforward consideration of the relationship between the momentum of a particle and the wavenumber of the corresponding matter wave, Hsu also concludes that Planck's constant $\hbar$ (strictly, Planck's constant divided by 2π), is not strictly a fundamental constant, but that the ratio $\hbar/c$, called J , should remain numerically invariant during transformation within the framework of proper time just as in special relativity proper. He points out that the value of the fine structure constant, whose value is $\bar{e}^2/J$, also remains a constant and thus retains in a system of common time its value of about $1/137$.

The significance of this argument will vary from one reader of it to another. Many will no doubt take the view that since the whole weight of experimental evidence supports standard special relativity, in which the velocity of light is strictly constant from one frame to another moving relative to it, it is a strictly academic exercise to attempt to distinguish between constants supposed on Hsu's argument to be composite and constants such as h/c which are strictly constant both in special relativity and in common time. Hsu's own justification is that, if the equations of motion are modified to allow for universal time, opportunities for using electromagnetic phenomena for telling the difference between one system and another are many fewer than would be expected. Hsu also argues that a system of common time would make intelligible concepts which are now hard to grapple with, such as that of a wave function purporting to describe the probability distribution of a particle at some instant of time.

A better way of tackling this problem may be that suggested by Panafiel and Rafanelli (*Il Nuovo Cimento* 72B, 157; 1982) who showed how it is possible to treat a classical system of interacting particles without violating Einstein's concept of time. □

Stellar rotation

Magnetic brakes on giant stars

from Robert Cannon Smith

ONE of the best-known properties of rotating stars is that cool stars rotate much more slowly than hot ones. The usual explanation has been twofold: gradual magnetic braking for cool main-sequence stars which have long had convective envelopes and so, presumably, solar-type winds; and evolutionary expansion coupled with conservation of angular momentum for cool giant stars, which have developed convective envelopes only recently. Precise testing of these explanations has been hampered by the difficulty of measuring low rotation speeds. But over the last few years it has become possible to measure rotation speeds to better than 1 km s^{-1} for bright stars and some very interesting new results are emerging that seem to show that the usual explanations are far from the whole story.

A recent series of papers¹⁻³ in *Astrophysical Journal* by David F. Gray of the University of Western Ontario has revealed a dramatic discontinuity in rotation speed at spectral-type G5 III which suggests that a magnetic brake also operates in giant stars, slowing them down much faster than the $t^{-1/2}$ rate found for cluster main-sequence stars by Skumanich⁴ and others. Gray also argues⁵ that main-sequence stars show evidence for an early period of rapid braking before settling into the $t^{-1/2}$ behaviour.

The observations behind these new results have become possible because of recent improvements in detectors which allow high signal-to-noise ratios to be obtained even at high spectral resolution. Gray used Reticon silicon diode arrays and coude spectrographs, making observations with the 2.1 m telescope at the McDonald Observatory of the University of Texas and with the 1.2 m telescope at the Elginfield Observatory of his own university. He obtained velocity resolutions of about 3 km s^{-1} and signal-to-noise ratios of better than 100.

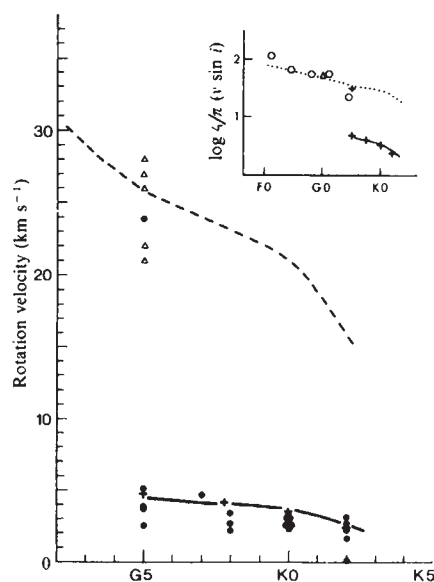
With data of that quality it is possible to use Fourier transform techniques^{6,7} to analyse the line profiles, which are Doppler-broadened by the variation of rotation velocity v over the stellar disc and by turbulence. In order to separate these effects, Gray assumes⁷ a microturbulence ξ with an isotropic gaussian velocity distribution and a macroturbulence ξ_{RT} described by a gaussian distribution with the restriction that the velocities are only radial or tangential to the stellar surface. He also makes the assumption, reasonable for slowly rotating stars, that there is no centre-to-limb variation in the intrinsic line profile, which is calculated from a grid of local thermodynamic equilibrium model atmospheres.

The results for giant stars are shown in

the figure. There is no way to determine the angle of inclination, i , of the rotation axis of a star to the line of sight, so what is plotted is the projected rotation speed $v \sin i$. The mean values (crosses) have been divided by the mean value of $\sin i$, $\pi/4$, to give an estimate of the actual rotation speed v for the cooler stars. The dramatic drop at G5 is obvious. (The only star that does not fit this pattern is ϕ Vir (not shown) with $v \sin i = 14.3 \text{ km s}^{-1}$ and spectral-type G2 III.) It seems that the surface layers at least, if not the whole star, have spun down from about 25 km s^{-1} to about 5 km s^{-1} in the time needed for a giant to evolve only a few tenths of a spectral class: less than a few million years.

The dashed line in the figure shows the effect of gradual evolutionary expansion with angular momentum redistributed within the star according to the prescription of Endal and Sofia⁸ but with total angular momentum conserved. The solid line shows the same prediction reduced by a factor of 5.7 to fit the drop at G5. The excellent fit to observations both before and after the discontinuity strongly suggests that the drop is caused by some braking mechanism that switches on suddenly, operates efficiently for a brief time and switches off as abruptly as it started.

A stellar wind carries away angular momentum and is a dramatically more effi-



Rotation velocity is shown as a function of spectral type. The closed circles are Gray's measured v and i values³ while the triangles are from a catalogue by A. Uesugi (Kyoto, 1976) and the open circles are mean values from ref. 16. The lines represent the behaviour expected from conservation of angular momentum (see text). (Copyright American Astronomical Society.)

cient brake if the star possesses a strong magnetic field that forces the wind to co-rotate with the star out to large distances. Gray therefore suggests that when the convective envelope of a giant star becomes deep enough, the interaction between rotation and convection produces dynamo action which creates a strong field. Once this field has reduced the rotation to a low level, via a stellar wind also supposed to be driven by motions in the convective envelope, the dynamo is no longer driven effectively, the field is reduced and the brake is released as sharply as it was applied.

Some supporting evidence for this picture comes from the X-ray emission from cool giants, which also drops abruptly at G5 III. In a recent preprint, Gray estimates that the ratio of X-ray to bolometric flux behaves like $F_X/F_{\text{bol}} = 3.6 \times 10^{-9} \langle v \rangle^2$ where $\langle v \rangle$ is the mean rotation speed (rather different from the linear angular velocity dependence found for RS CVn systems⁹). The X-ray emission is taken to be an indicator of surface magnetic field strength, so the drop is consistent with a dynamo's suddenly turning off.

However, solar X rays come from regions of closed magnetic loops¹⁰, while the solar wind escapes along open field lines through coronal holes where there is no X-ray emission. There is therefore no guarantee that strong X-ray emission means efficient magnetic braking: if X-ray emission implies closed field lines and a weak wind, it may mean the exact opposite. Furthermore, other indicators of chromospheric activity, such as the CIV flux, show no clear correlation with the rotation of these giant stars. Nonetheless, the idea that magnetic braking is involved is attractive and plausible and the drop in surface activity could perhaps signal the onset of a strong stellar wind and so a transition to a phase of rapid braking¹¹.

A by-product of these rotation measurements is a set of values³ for the micro- and macroturbulence in giant stars. If these values represent real turbulent motions, then we can begin to learn something about how turbulence varies with temperature and depth in cool giant envelopes. The microturbulence is poorly determined but is consistently larger for the strong lines, which are formed further out in the photosphere than the weak lines. This could suggest³ wave motions whose amplitude increases with the outward decrease in density. The microturbulence decreases steadily with effective temperature T_{eff} , $\xi_{\text{RT}} - T_{\text{eff}}^{2.6}$, at least for the weak lines; the strong lines show a more complicated behaviour. Non-radial oscillations are predicted to give a $T^3 g^{-1/4}$ dependence, which is a reasonable fit to late-G and early-K stars¹².

Accurate rotation velocities are also becoming available for slowly rotating main-sequence stars¹³, but here the picture is much less clear. Gray argues⁵, by analogy with the giant stars, that all stars just arriving on the main sequence undergo a period of rapid braking before settling down to a

slow loss of angular momentum of the solar type. By using the upper envelope of the observed distribution of rotation speed with spectral type and comparing it with the angular momentum-mass distribution for upper-main-sequence stars he claims that the rapid braking ceased abruptly at a rotation speed that was the same for all stars of the same spectral type and was about 2–2.5 times the present maximum observed values.

Some models of solar-wind braking do show a rapid initial braking¹⁴ if the magnetic flux is proportional to the angular velocity, but the timescale is a few times 10^8 years, whereas the giant stars must be braked within a few times 10^6 years. Since there is no direct evidence for a phase of very rapid braking for main-sequence stars, and there are many theoretical uncertainties about the approach to the main sequence, any clear picture of the early rotational history of main-sequence stars must await more data on very young main-sequence stars. For the

moment, the observations are qualitatively consistent with a solar-type loss of angular momentum, with the production of more effective winds by the deeper convective envelopes in cooler stars, and the loss of much of the angular momentum in the pre-main-sequence phase¹⁵. □

Robert Connon Smith is Lecturer in Astronomy in the Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH.

1. Gray, D.F. *Astrophys. J.* **251**, 155 (1981).
2. Gray, D.F. *Astrophys. J.* **254**, 162 (1982).
3. Gray, D.F. *Astrophys. J.* **262**, 682 (1982).
4. Skumanich, A. *Astrophys. J.* **171**, 565 (1972).
5. Gray, D.F. *Astrophys. J.* **261**, 259 (1982).
6. Carroll, J.A. *Mon. Not. R. astr. Soc.* **93**, 478 (1933).
7. Gray, D.F. *The Observation and Analysis of Stellar Photospheres* (Wiley, New York, 1975).
8. Endal, A.S. & Sofia, S. *Astrophys. J.* **245**, 671 (1981).
9. Walter, F.M. & Bowyer, S. *Astrophys. J.* **245**, 671 (1981).
10. Svalgaard, L. & Wilcox, J.M. *A. Rev. Astr. Astrophys.* **16**, 429 (1978).
11. Durney, B.R., Mihalas, D. & Robinson, R.D. *Publs astr. Soc. Pacif.* **93**, 537 (1981).
12. Smith, M.A. *Astrophys. J. Lett.* **242**, L115 (1980).
13. Soderblom, D.R. *Astrophys. J.* **263**, 239 (1982).
14. Belcher, J.W. & MacGregor, K.B. *Astrophys. J.* **210**, 498 (1976).
15. Vogel, S.N. & Kuhi, L.V. *Astrophys. J.* **245**, 960 (1981).
16. Fukuda, J. *Publs astr. Soc. Pacif.* **94**, 271 (1982).

Biophysics

Muscle crossbridges in disarray

from Robert A. Mendelson

MUSCLE contraction has been thought to occur by the asynchronous rocking of myosin crossbridges between the parallel arrays of thick myosin filaments and thin actin filaments to produce a longitudinal translation between the two. This has been presumed, largely on the basis of X-ray diffraction and electron microscopy, to occur between the 45° orientation seen in rigor muscle and the relaxed orientation, in which nearly all the crossbridges have been thought to be roughly normal to the fibre axis with their distal ends towards the thin-filament-binding site. In this issue of *Nature*, however, Poulsen and Lowy¹ report that a significant fraction of myosin crossbridges in relaxed vertebrate muscle scatter X rays as if they were in solution rather than locked into the thick filament lattice.

The main-line technique for the study of filament movements, X-ray diffraction from living fibres, has been focused on the changes in the reflections arising from thin (mostly actin) and thick (mostly myosin) filaments. Isometrically contracting fibres have failed to show any new reflections or any significant change in intensity on the actin-based layer lines². Myosin reflections, except the 14.3 nm meridional, only decrease in intensity. Although the interpretation of these patterns is unclear, it appears that at least azimuthal disordering is required. The difficulties in interpretation of thick filament patterns and their changes may be relieved in the future by the work of Poulsen and Lowy, who observe that the diffuse scattering seen in relaxed fibres originates largely from randomized

myosin crossbridges (S1) rather than other proteins, and thus diffuse scattering data must also be considered in models of crossbridge action.

Poulsen and Lowy find that the shape of the diffuse portion of the scattering curve is circularly symmetric and its intensity profile matches that from isolated S1-solution scattering data. This match occurs within an angular range whose lower bound is large enough to eliminate particle-particle interference and excluded-volume effects, and whose upper bound minimizes the influence of S2 cross-sectional scatter on the scattering curve. When a relaxed fibre is stretched to a length where thick and thin filaments no longer overlap, there is a slight increase in disordered intensity. However, if non-overlapping muscle is then put into rigor a further increase in intensity of the disordered fraction occurs and almost no reflections can be seen. The authors take this situation, in which the thick filament structure is almost certainly perturbed, as being the base-line fully disordered intensity.

Modelling the decrease in the diffuse scattering curve at small angles relative to S1 solution data, the authors find that an 8 nm root-mean-square displacement (in every direction) is required. However, the layer-line intensity does not fall off rapidly enough with increasing angle to make such single-state crossbridge models tenable. In a model in which one fraction is completely ordered, the second fraction should have 9 nm displacement. The second fraction would be about 45 per cent in stretched muscle, and extrapolation of the authors'

data indicates it would be about 30 per cent in fully overlapped muscle.

On passing into rigor at full overlap, the diffuse intensity drops by a factor of about 1.8 and the pattern becomes compressed diagonally. This remaining intensity arises from crossbridges occupying actin sites randomly and thus creating substitution disorder. Modelling studies show that the observed compression can only be created by a crossbridge having a bend of about 45° .

How do these studies correlate with those using different methods? Fluorescence depolarization³ and saturation transfer electron paramagnetic resonance⁴ solution experiments, in which dipolar probes are rigidly attached to a single fast-reacting S1 thiol, indicate that brownian rotational movement of S1 moieties is much greater when myosin is in monomeric (soluble) form than when part of a synthetic or native thick filament. Presumably this large decrease in mobility is caused either by steric hindrance of the partner head or by the thick filament backbone, or both.

These methods have so far failed to detect a significant component of free or nearly free S1 motion in thick filaments of myofibrils. On the other hand, experiments aimed at measuring the average orientation of these probes in oriented muscle fibres⁵ have been unable to detect any ordering. Thus, probe results give a picture of scrambled heads with little mobility, presumably because at least a part of S1 is near the thick filament backbone. Since it is known that glycerinated rabbit *psaos* fibres have less order than living frog muscle, the lack of order found in these experiments perhaps should not be compared with the X-ray results. It is however unclear whether the model of the disordered fraction used by Poulsen and Lowy is consistent with the probe results. Fortunately, further probe⁶ and X-ray⁷ work is under way to study this question.

The current X-ray results raise fresh questions about the functional significance, if any, of the relaxed crossbridge order. For example, must myosin heads be ordered or disordered for attachment to the thin filament upon activation? Are myosin heads briefly adopting one of these states in active muscle? Whatever the significance of these results it is clear that simultaneous studies of both crossbridge order and orientation will be useful, if not obligatory, for unravelling the binding and micromechanism of crossbridges in contracting muscle. □

Robert A. Mendelson is at the Department of Biochemistry and Biophysics, University of California, San Francisco, California.

1. Poulsen, F.R. & Lowy, J. *Nature* **303**, 146 (1983).
2. Huxley, H.E. *et al. J. molec. Biol.* **158**, 637 (1982).
3. Mendelson, R.A. & Cheung, P. *Science* **194**, 190 (1976).
4. Thomas, D., *et al. Biophys. J.* **32**, 873 (1980).
5. Thomas, D. & Cooke, R. *Biophys. J.* **32**, 891 (1980).
6. Wilson, M.G.A. & Mendelson, R.A. *J. Muscle Res. Cell Motil.* **3**, 464 (1983).
7. Kress, M. *J. Muscle Res. Cell Motil.* **3**, 465 (1983).

Plant ecology

Revival of the organismal heresy

from Peter D. Moore

THE early decades of this century found plant ecologists embroiled in largely semantic arguments concerning the nature of plant communities. Views tended to polarize between those of Clements¹, who was impressed by the integration of component species and regarded the community as analogous to an organism; and Gleason², who developed an individualistic model in which most species behaved independently of one another and the community was therefore a fortuitous juxtaposition of plant individuals, its composition being limited only by the environmental amplitudes of its members. The debate has largely lost its steam, though the philosophical divide remains. Each side now seems prepared to tolerate, and even respect, the other's viewpoint.

It is strange then to find that the 'organismal heresy' is once again being raised — this time the apparent convergent evolution of geographically separated ecosystems experiencing similar stimuli. The origin of this renewal and extension of the old debate is to be found in Australia*, and in one Australian ecosystem in particular — heathland. This is a sclerophyllous shrub-dominated habitat abundant through much of the Australian continent. Its similarities to heathland ecosystems elsewhere in the world and whether it is per-

missible to speak of the evolution of entire ecosystems are the subjects of debate.

Specht³ considers that all the heathlands of the world have three features in common: (1) dominance by evergreen sclerophyllous plants, (2) the presence of (but not necessarily dominance by) plants of the Ericales and (3) confinement to soils poor in plant nutrients. There is certainly a general resemblance, at least in physiognomic terms, between the heathlands of, for example, north-west Europe, the Mediterranean, South Africa and Australia, but there are also some contrasting features.

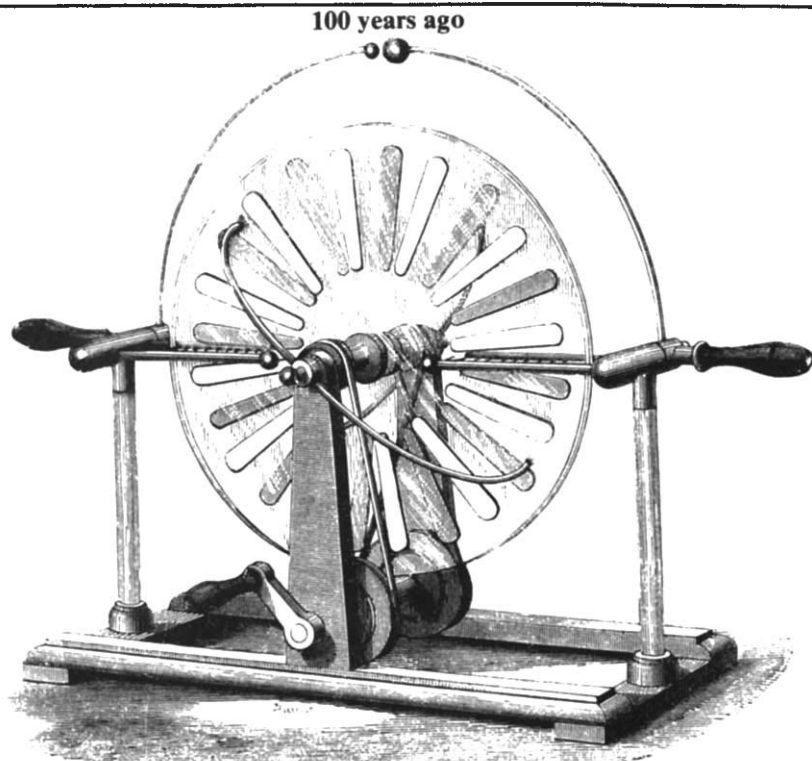
One notable characteristic of the northern European heathlands⁴ is that most of the important plant components have small light seeds easily dispersed by wind. This may well reflect the fact that heathlands in this part of the world have developed relatively recently (during the past 10,000 years) as a result of human deforestation⁵ and have been maintained by fire. They have therefore developed and regenerated in a patchwork mosaic, within which a capacity for long-distance seed dispersal has been advantageous.

In Southern Hemisphere heaths, seed strategies are more variable. Often prominent are members of the Proteaceae — a family in which some members have small

wind-dispersed seeds and others have large seeds. In a recent survey of proteaceous seeds, Kuo, Hocking and Pate⁶ found that the large seeds were generally very rich in plant nutrients, with the cotyledons acting as major storage organs. Often the levels of phosphorus and nitrogen in these tissues were 30–500 times and 8–100 times respectively the levels of these elements in other parts of the mature plant, such as the leaves. Hocking⁷ has observed the accumulation of nutrients in the fruits of the proteaceous genera *Grevillea* and *Hakea* and concluded that the nutrient build-up represents a strategy which is of special value in impoverished soils. Presumably the developing seedling depends on these reserves for some period of time during its establishment.

Milewski⁸ has pointed out that the Australian heathlands are particularly poor in potassium and, if South Africa and Australia are compared, soft-bodied fruits, attractive to vertebrates, are far commoner in South African heathland species, especially in areas with a better supply of potassium. Seed dispersal in the large-seed Australian heath plants, he claims, is more frequently performed by ants than by vertebrates, and could be associated with a lack of rain forest species in the flora and hence fewer fleshy-fruited bird-dispersed types in Australia. Rain forest plant families, such as Apocynaceae, Oleaceae and Tiliaceae, occur down the eastern side and on the southern tip of Africa, but in Australia they are restricted to the north-east. Although this could have limited their contribution to the Australian heathland ecosystems, Milewski considers that the lack of potassium in the soils is a more important factor.

The study of the evolution of ecosystems requires a knowledge of past history if developmental parallels are to be clarified. In Australia, palynological work has provided evidence of widespread wet evergreen rain forest during the early Tertiary which was increasingly replaced by drier open vegetation especially in Miocene–Pliocene times⁹. Casuarinaceae and Proteaceae are well represented in sediments of Miocene age and there are high levels of *Eucalyptus* pollen from the



Wimshurst's influence machine. "Theoretically, a small initial charge must be imparted to one or more of the carriers or to one of the two main conductors. Practically, if dry and free from dust, the machine excites itself, and, after a couple of turns have been given to the handle, discharges sparks freely." From *Nature* vol. XXVIII, 13 (1883).

*An Ecological Society of Australia symposium entitled 'Are Australian ecosystems different?' will be held in 1984.

1. Clements, F.E. *Plant Succession. An Analysis of the Development of Vegetation* (Carnegie Instn, Washington, 1916).
2. Gleason, H.A. *Bull. Torrey bot. Club* 53, 7 (1926).
3. Specht, R.L. *Heathlands and Related Shrublands* (Vol. 9A of *Ecosystems of the World*) 1 (Elsevier, Amsterdam, 1979).
4. Gimingham, C.H. *Ecology of Heathlands*, 263 (Chapman & Hall, London, 1972).
5. Dimbleby, G.W. *Oxford For. Mem.* 23, 1 (1965).
6. Kuo, J., Hocking, P.J. & Pate, J.S. *Aust. J. Bot.* 30, 231 (1982).
7. Hocking, P.J. *Aust. J. Bot.* 30, 219 (1982).
8. Milewski, A.V. *J. Biogeogr.* 9, 505 (1982).
9. Galloway, R.W. & Kemp, E.M. in *Ecological Biogeography of Australia* (ed. Keast, A.) (Junk, The Hague, 1981).
10. Sluiter, I.R. & Kershaw, A.P. *Alcheringa* 6, 211 (1982).
11. Kershaw, A.P. & Sluiter, I.R. *Aust. J. Bot.* 30, 279 (1982).
12. West, R.G. *The Pre-glacial Pleistocene of the Norfolk and Suffolk Coasts* (Cambridge University Press, 1980).

Pliocene in Western Australia¹⁰, so the open scrub heath can be assumed to have been present at this time, while local survival of the rain forest continued in the north-east¹¹. Undoubtedly the heathland has been further favoured in the late Quaternary by the arrival of man and an associated increase in the frequency of fire.

There is, therefore, a considerable difference in age between the Australian and the northern European heathlands. Although there are records of heathland pollen assemblages from pre-glacial Pleistocene sediments in Britain¹², such ecosystems have subsequently been

regularly disturbed by glacial episodes. In Australia, the Pliocene heaths appear to have had a long period in which to evolve without interruption by such extreme factors as glaciation. This may account for some of the observed differences between Northern and Southern Hemisphere heaths in terms of the evolution of seed strategies. Convergent evolution might well continue, were ours given sufficient time to catch up. □

Peter D. Moore is at the Department of Plant Sciences, University of London King's College, 68 Half Moon Lane, London SE24 9JF.

Cell signalling

Dual control of adenylate cyclase

from Miles Houslay

MANY hormones exert their effects on target cells by elevating the intracellular concentration of cyclic AMP, which acts as a 'second messenger'. They achieve this by binding to specific receptors on the cell surface, which then activate the enzyme adenylate cyclase. The interaction between these two entities is mediated by a distinct regulatory protein (N_i) which is dependent on guanine nucleotides for function. Recently, excitement has focused on observations that certain receptors, for example muscarinic, opiate and α -adrenoceptors, actually lead to inhibition of adenylate cyclase activity. Indeed, it would seem that in many cells, adenylate cyclase is under dual control. Furthermore, guanine nucleotides are essential for the expression of these inhibitory effects. This has raised the question of whether a single guanine nucleotide regulatory protein mediates both stimulatory and inhibitory effects, or whether distinct regulatory proteins exist. Very recent observations from several laboratories¹⁻⁶ have suggested that the latter alternative is correct. In this edition of *Nature*⁷, Jakobs and co-workers demonstrate that a fully functional hormone- and guanine-nucleotide-controlled inhibitory response is displayed by a somatic cell variant (cyc^-) which is genetically deficient in the stimulatory guanine nucleotide regulatory protein, N_s .

The S49 mouse lymphoma cell line has proved, yet again, to be a powerful tool for identifying the molecular components of the adenylate cyclase system. This cell possesses a β -adrenoceptor that stimulates adenylate cyclase activity in the presence of β -agonists. However, high concentrations of cyclic AMP are toxic to the cells, enabling mutants defective in cyclic AMP metabolism to be selected. One such mutant, cyc^- , has proved invaluable⁸. Although it possesses a β -adrenoceptor and catalytic unit of adenylate cyclase, β -agonists do not stimulate cyclic AMP production, and guanine nucleotides, sodium fluoride and cholera toxin also fail to

stimulate the enzyme. Indeed, the activity of adenylate cyclase is so low that originally cyc^- was thought to be deficient in this enzyme. Later studies showed, however, that it could be exposed if Mn^{2+} rather than Mg^{2+} was used as co-substrate and in a key experiment, Ross and Gilman⁹ demonstrated that both hormone- and guanine-nucleotide-stimulated activities could be reconstituted in cyc^- membranes using a detergent extract from the plasma membranes of wild-type cells whose adenylate cyclase had been heat-inactivated. This led to the suggestion that what was missing from cyc^- cells was a distinct guanine nucleotide regulatory protein, N_i ; and this protein has since been purified to homogeneity from rabbit liver¹⁰.

The same cyc^- variant has now been used by groups in Cr teil, Houston and Heidelberg^{5,11} to identify a distinct (N_i) guanine nucleotide regulatory protein that mediates inhibitory responses. To detect inhibition of the adenylate cyclase in cyc^- membranes it was necessary first to amplify the residual activity of the enzyme. Mn^{2+} could not be used as it not only perturbs guanine nucleotide responses but it also 'uncouples' both stimulatory and inhibitory effects. The trick proved to be the use of forskolin, a diterpene which can act as an extremely powerful activator of adenylate cyclase¹², which made it possible to show that the adenylate cyclase activity of cyc^- cells was indeed inhibited by guanine nucleotides. Such inhibition was elicited using much lower concentrations of non-hydrolysable analogue of GTP than with GTP itself, whereas the opposite is true for stimulatory responses. With Jakobs and co-workers' observation⁷ that somatostatin produces a guanine nucleotide-dependent inhibition of adenylate cyclase here, it seems that a complete hormone-controlled inhibitory system exists in cyc^- cells.

The nature and mechanism of action of this inhibitory regulatory protein now re-

quires elucidation. A useful tool may prove to be islet-activating protein (IAP) from *Bordetella pertussis*. IAP blocks the effect of inhibitory hormones on adenylate cyclase, apparently through the ADP-ribosylation of a distinct membrane protein¹³. This target protein has a molecular weight of 39–41,000 and is not a substrate for cholera toxin, which specifically ribosylates subunits of molecular weight 52,000 and 45,000 forming part of N_s . The IAP target protein has been purified and seems to be a heterodimer consisting of both 41,000- and 35,000-molecular-weight polypeptides¹⁴. It has strikingly similar physical properties to N_i and may well turn out to be N_i itself. Indeed, Birnbaumer and co-workers⁶ have very recently demonstrated a 39,000-molecular-weight substrate for IAP in membranes from cyc^- cells. Moreover, treatment with IAP renders the forskolin-stimulated adenylate cyclase insensitive to inhibition by GTP. However, IAP does not block the inhibitory effects of a non-hydrolysable analogue of GTP on cyc^- adenylate cyclase, which suggests that the system has a considerable complexity. As to mechanism, it has been suggested that the regulatory proteins are both GTPases with the GTP form being the active state; this suggestion has however been criticized, not least because purified functional N_i does not exhibit GTPase activity.

Whether further distinct guanine nucleotide proteins linked to specific receptors will be discovered, as Rodbell¹⁵ has suggested, remains to be seen. Indeed, in this light there is the recent provocative suggestion^{16,17} that the insulin receptor, whose mechanism has defied comprehension for many years, may exert some of its effects through interaction with a specific guanine nucleotide regulatory protein with properties somewhat similar to those of N_i .

Miles D. Houslay is Reader in Biochemistry at the University of Manchester Institute of Science and Technology, Manchester M60 1QD.

1. Stiles, G.L. & Lefkowitz, R.J. *J. biol. Chem.* **257**, 6287 (1982).
2. Insel, P.A., Stengel, D., Ferry, N. & Hanoune, J. *J. biol. Chem.* **256**, 7485 (1982).
3. Motulsky, H.J. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 4193 (1982).
4. Seamon, K.B. & Daly, J.W. *J. biol. Chem.* **257**, 11591 (1982).
5. Hildebrandt, J.D., Hanoune, J. & Birnbaumer, L. *J. biol. Chem.* **257**, 14723 (1982).
6. Hildebrandt, J.D. *et al. Nature* **302**, 706 (1983).
7. Jakobs, K.H., Aktories, K. & Schultz, G. *Nature* **303**, 177 (1983).
8. Bourne, H.R., Coffino, P. & Tomkins, G. *Science* **187**, 750 (1975).
9. Ross, E.M. & Gilman, A.G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3715 (1977).
10. Sternweis, P.C., Northup, J.K., Smigel, M.D. & Gilman, A.G. *J. biol. Chem.* **256**, 11517 (1981).
11. Jakobs, K.H., Gehring, U., Gaugler, T., Pfeuffer, T. & Schultz, G. *Eur. J. Biochem.* (in the press).
12. Seamon, K.B. & Daly, J.W. *Trends pharmac. Sci.* **4**, 120 (1983).
13. Katada, T. & Ui, M. *J. biol. Chem.* **257**, 7210 (1982).
14. Bokoch, G.M., Katada, T., Northup, J.K., Hewlett, E.L. & Gilman, A.G. *J. biol. Chem.* **258**, 2072 (1983).
15. Rodbell, M. *Nature* **284**, 17 (1980).
16. Heyworth, C.M., Rawal, S. & Houslay, M.D. *FEBS Lett.* **154**, 87 (1983).
17. Heyworth, C.M., Wallace, A.V. & Houslay, M.D. *Biochem. J.* (in the press).

Genes and development

Cloning the genes that specify fruit flies

from Geoffrey North

AFTER the hundred years or so of the science of developmental biology, it can safely be said that we still have little idea how the linear array of nucleotides along a DNA chain can direct the production of a complex multicellular organism. However, at a recent meeting of the British Society for Cell Biology* classical and molecular geneticists came together to discuss what are surely the first steps in solving this seemingly intractable problem. Not surprisingly, the organism that is beginning to provide the answers is the fruit fly, *Drosophila melanogaster*.

Since *Drosophila* genetics began, around the beginning of this century, a number of regions of the genome have attracted particular attention because of the remarkable effects on development of mutations mapping within them. Some cause the replacement of one body structure by another; others cause the complete absence of particular parts of the body. These regions of the genome are the obvious targets of molecular biologists interested in development.

Some of these regions have now been cloned, and the task of taking them apart nucleotide by nucleotide has begun. Already, a complexity fully commensurate with that of their classical genetics has been revealed. Giant transcription units, processed into multiple discrete small RNA products in a way that changes during development, appear to be the rule.

Specification of the body pattern

The segmented body pattern of an adult fruit fly (Fig.1) is established early in development by the setting-aside of groups of cells destined to give rise to specific structures in the adult¹. The fate of each group of cells appears to be regulated by a set of genes, mutations in which can lead to the replacement of one segment by another, or of one part of a segment by the homologous part of another segment (see,

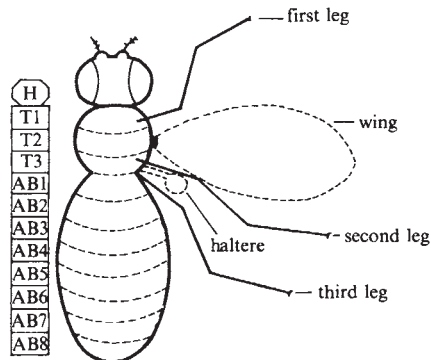


Fig. 1 Segmental structure of an adult fruit fly. The segments are indicated on the left: H, head segment(s); T1-3, thoracic segments (T1, prothoracic; T2, mesothoracic; T3, metathoracic); AB1-8, abdominal segments.

for example, Fig.2).

These are the so-called homoeotic genes, most of which cluster in two regions of the genome: the *bithorax* complex, whose products influence the development of segments of the fly posterior to thoracic segment 2; and the *Antennapedia* complex, whose products contribute to the specification of anterior structures including the head.

The genetics of the *bithorax* complex have been worked out mainly by Edward Lewis of the California Institute of Technology². Recessive mutations of the *bithorax* complex tend to cause the replacement of one segment by a more anterior one and dominant mutations, conversely, the replacement of anterior structures by more posterior ones; and remarkably, the genetic map of the mutations is roughly colinear with the anterior-posterior locations of the segments affected (see Fig. 3). Lewis has speculated that this pattern of developmental control may directly reflect the evolution of the fly from a more primitive, uniformly segmented worm-like creature — the second thoracic segment representing the prototype segment, on which the products of successive *bithorax* components act to produce the characteristics of the successively more posterior segments. This idea has been further developed by Peter Lawrence (MRC, Cambridge), Antonio Garcia-Bellido and Gines Morata (Madrid University), who have formulated the notion of a 'genetic address' consisting of a specific pattern of expression of different combinations of homoeotic genes defining the developmental fate of different groups of cells early in development¹.

This notion of a developmental address is elegantly supported by some recent studies on the *Antennapedia* complex (Fig. 4). The classical *Antennapedia* mutation is a dominant mutation of the *Antp* locus that causes the replacement of antennae by second-thoracic legs. The dominance of this mutation implies that it results in inappropriate expression of the *Antp* gene in the head segment, and Gary Struhl (Harvard University) has suggested that the normal function of the *Antp* gene is to prevent the genes involved in making antennae in the head from acting incorrectly in other segments of the body³. Thus in *Antp* mutants, in which the *Antp* gene is removed altogether, antennae are formed where one would normally expect second-thoracic legs. The *Antp* mutation on its own, however, is not sufficient to produce antennae in place of first- and third-thoracic legs. These transformations require other mutations that on their own would transform first- or third- into

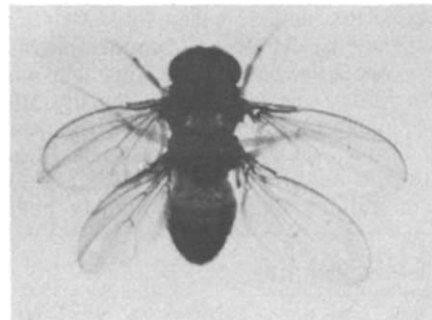


Fig.2 A four-winged adult fruit fly showing a complete transformation of the third-thoracic into a second-thoracic segment. The fly is a hemizygote, with one chromosome carrying a deletion of the *bithorax* complex, the other carrying *bithorax* (*abx*, *bx*³) and *postbithorax* (*pbx*) mutations. (Photograph by kind courtesy of Edward Lewis, California Institute of Technology and Michael Akam, University of Cambridge.)

second-thoracic legs. This is precisely consistent with the idea of a genetic address, the two mutations being required to spell the correct code word for antenna construction.

Although Lewis found a neat correspondence between the relative positions of mutations on the genetic map and the relative positions of the structures affected by the mutations, any idea that the *bithorax* complex is an array of simple genes is untenable. Morata described how on the basis of complementation studies the various components of the *bithorax* complex are apparently grouped into three functional units. Within each unit, interactions seem to occur along the chromosome (that is, in *cis*); for example, *Ubx* mutations seem to inactivate the entire region of the left hand of the *bithorax* complex from *Ubx* to *pbx* (see Fig.3) so that, for instance, *Ubx/bxd* flies have the same phenotype as *bxd/bxd* flies. This complexity is reflected in the first fruits of the molecular analysis of these regions of the genome, which are described below.

The *bithorax* complex

How does one clone a region of the genome without an identified product? A different approach is clearly needed from that used to clone, say, the β -globin gene. Several solutions to this problem are illustrated in the examples that follow, and again it will be seen how the classical genetics and cytogenetics of *Drosophila* have proved crucial.

The approach of Welcome Bender and Pierre Spierer, who set out to clone the *bithorax* complex in the laboratory of David Hogness at Stanford University over four years ago, was that of 'chromosome-walking' from the closest gene to the *bithorax* complex that had already been

Fig.3 Genetic map of the *bithorax* complex, modified from ref.2. The left-hand *Ubx* site has recently been mapped by Michael Akam (University of Cambridge). The map shows the relative positions of point mutations that cause the following phenotypes: *Ubx* (*Ultrabithorax*), a conversion of the third-thoracic and first-abdominal into second-thoracic segments; *bx* (*bithorax*), conversion of the anterior compartment of the third-thoracic into a second-thoracic segment; *Cbx* (*Contrabithorax*), a dominant mutation converting the posterior second-thoracic into posterior third-thoracic segment; *bxd* (*bithoraxoid*), converts first-abdominal into a third-thoracic segment; *pbx* (*postbithorax*), like *bx*, but for the posterior instead of the anterior compartment; *iab-2* to *iab-8* (*infra-abdominal*) cause the conversion of abdominal segments into more anterior abdominal segments. The DNA scale is marked in kilobases from an arbitrary zero at the site of the initial bridgehead in the *bithorax* complex. <-----> Indicates the region within which map chromosome breakpoints in rearrangements causing the marked phenotypes. The arrowhead on the high-molecular-weight transcripts indicates the direction of transcription (5'-3'), and the hatching indicates the approximate positions of the major *Ubx* transcripts. The *Ubx* transcription data are based principally on the work of Robert Saint, and the *bxd* data on that of Michel Goldschmidt-Clermont, both members of the Hogness group at Stanford University.

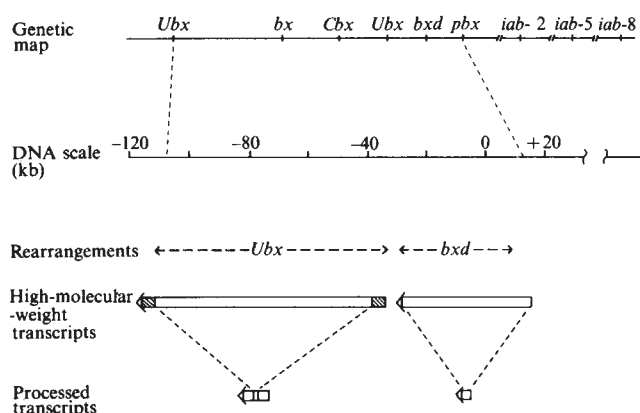
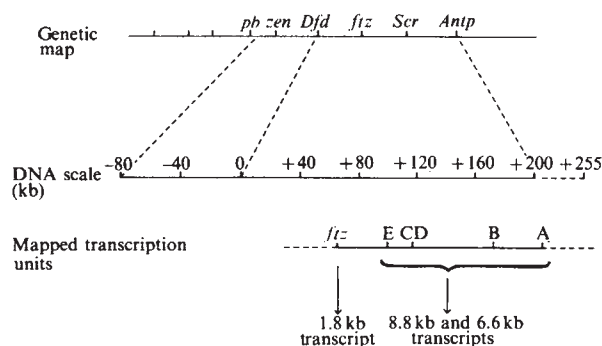


Fig.4 Genetic map of the *Antennapedia* complex. The phenotypes of the mutations of the loci marked are: *pb* (*proboscipedia*), conversion of the mouth-parts (labial palps) into either antennae or legs; *zen* (*zerknüllt*), defective in gastrulation; *Dfd* (*Deformed*), disruption of the eye tissue; *ftz* (*fushitarazu*), affects the number of segments formed; *Scr*, converts the first-thoracic into a second-thoracic segment; *Antp* (*Antennapedia*), dominant mutations convert antennae into legs, recessive mutations convert legs into antennae. The DNA scale delineates the region cloned so far by Kaufman's group (Scott, University of Indiana) in a chromosome-walk that began at the *Dfd* locus. E, CD, B and A mark the regions of the *Antp* transcription unit apparently encoding the exons of the 8.8 kb and 6.6 kb transcripts. This transcription unit is at least 105 kb long.



cloned. This is a rather laborious process, involving the identification of overlapping fragments of DNA in a cloned genome library, and inching along the chromosome, checking your progress *en route* by *in situ* hybridization to polytene chromosomes. Starting as they did over 4,000 kilobases (kb) away from the *bithorax* complex, it would have taken Bender and Spierer many years to reach their destination; however, classical genetics came to their aid with an inversion mutation which allowed them to 'jump' a large stretch of the chromosome and continue their walk from a new bridgehead within the *bithorax* complex itself.

Bender has now walked over 300 kb from the original bridgehead, and, as Michael Akam (University of Cambridge) reported, the use of this cloned DNA has already produced some startling results. The region studied in most detail so far is that towards the left of the genetic map shown in Fig.3, from the left-hand site of *Ubx* point mutations to the site of *postbithorax* (*pbx*) mutations, which seem to form a single complementation group. Two major transcription units have been identified in this region, both of which are processed to give multiple products. The *Ubx* transcripts first appear at the blastoderm stage of embryogenesis, and the particular pattern of discrete processed RNAs changes during development.

A 3.0 kb *Ubx* transcript has been studied in most detail. It is constructed from two major exons, which are a remarkable 70 kb

apart in the genome. Sequencing of a cDNA derived from this RNA has shown the two major exons to be separated by a 102 base pair sequence constructed from two 'micro-exons'. (David Hogness reported at a recent conference in Cambridge[†] that a second 3.0kb *Ubx* cDNA appears to have a different sequence to that of the first — though they share some exons — so the transcript pattern may be even more complex than at first sight.) These results fit neatly with the pattern of mutations: the two widely separated sites of *Ubx* point mutations map within the two major exons, and the entire 75 kb transcription unit delineates the region of the chromosome in which map the breakpoints of chromosome rearrangements that also cause the *Ubx* mutant phenotype, presumably by irrevocably separating the two exons. Other transcripts of different sizes have been identified sharing the same 5' exon as the 3 kb RNA, but with different 3' exons. It may be that these transcripts include those affected by *bithorax* (*bx*) mutations, which map within the 70 kb intervening sequence and cause only a partial *Ubx* phenotype (only the anterior compartment of the third-thoracic segment is affected).

The availability of cloned DNA derived from the *Ubx* 5' exon has made it possible to see which cells are expressing the *Ubx* transcription unit at different stages of development. Akam described how sections of entire embryos or larvae could be prepared so that the RNA in each cell is accessible for hybridization to a complemen-

tary radioactively labelled DNA probe. This is the technique of RNA *in situ* hybridization, which has also been used to study *Antennapedia* expression (see later). In embryos, the highest levels of *Ubx* expression are in the cells of the ventral ganglia of the posterior segments, whereas no *Ubx* expression is found in the most anterior ganglia — as would have been predicted from the effects of mutations.

The imaginal discs of larvae, which evert during metamorphosis to give rise to the cuticular structures of the adult, also give results clearly consistent with the genetics: no expression in the discs destined to make the wings and second legs (second thoracic), most in the ones that will make the halteres and third legs (third thoracic). (Most hybridization is observed clustered over nuclei, which may reflect the great length of the transcription unit — at typical rates, it would take 45 min to make a single transcript — resulting in a large number of nascent transcripts anchored to the chromosome by RNA polymerase at any one time.)

The *Antennapedia* complex

The *Antennapedia* complex has been cloned by two groups, Tom Kaufman's at University of Indiana (represented by Matthew Scott) and a group at the Biozentrum in Basel (represented by its leader Walter Gehring). Scott explained how one approach used in the cloning of the *Antennapedia* complex was literally to cut it out of a giant polytene chromosome, the banding pattern of which clearly tells you

where to cut (actually carried out by V. Pirrotta and F. Scalenghe at Heidelberg). However, chromosome-walking was again the most important technique, which they carried out both from a nearby α -tubulin gene and from a bit of DNA that had jumped into the middle of the *Antennapedia* complex of a *Deformed* mutant, part of which had already been cloned by Bender. The extent of the cloning by the Indiana group is shown in Fig. 4.

In the region of the *Antp* locus, a huge transcription unit of at least 105 kb has been identified (Scott; Gehring), the primary product of which is processed into two major small RNAs, both of which first appear in embryogenesis at the blastoderm stage. One, 6.6 kb long, disappears around the time of segmentation; the other, 8.8 kb long, is present until the end of embryogenesis, when it disappears only to reappear again in late (third-instar) larvae.

As described above, the classical *Antp* mutation is dominant, apparently causing inappropriate *Antp* expression in the head segments. It turns out that nearly all such mutations are chromosomal rearrangements, and seem to involve a special class of breaks that map within a 70 kb region of the *Antp* transcription unit.

The question of how the resulting disruption of the normal transcription unit could lead to a gain-of-function mutation may be answered by RNA *in situ* hybridization experiments⁴ analogous to those on *bithorax*, and reported by Gehring. Like *Ubx*, *Antp* seems to be expressed early in development in cells destined to give rise to the larval and adult nervous system. At the larval stage most *Antp* RNA is found in the wing (second-thoracic) discs, and some in the antennal discs. The obvious next experiment will be to see whether inappropriate *Antp* expression can be seen in the antennal discs of dominant *Antp* mutants.

Included in the region of the *Antennapedia* complex that Kaufman's group have so far cloned is an interesting non-homoeotic gene called *fushitarazu* (*ftz*, see Fig. 4). Mutations in this gene affect the actual number of segments formed, and so it may be an example of a class of genes Tom Kornberg (University of California, San Francisco) has termed 'merismatic' (from the Greek *merismos*, meaning to subdivide the whole into parts), which direct the process by which groups of cells with different developmental fates are set aside during development. A discrete *ftz* transcript has been identified which is expressed early in embryogenesis, at the time defined by experiments with temperature-sensitive *ftz* mutants to be that at which wild-type *ftz* gene product is required for normal development.

Engrailed and Scute/Achaete genes

Other genes affecting development fall outside either the *bithorax* or the *Antennapedia* complex. The *engrailed* gene, for example, has properties characteristic of both a homoeotic and merismatic gene. It

may be considered merismatic because it is involved in the subdivision of cells within each segment into two groups, which are destined to become the anterior and posterior compartments of each segment¹. Expression of the *engrailed* gene in posterior cells appears to prevent them from mixing with anterior cells. It is also a homoeotic gene, because it forms part of the genetic code word directing posterior cells to produce posterior structures, and not anterior structures.

Kornberg reported that his and Patrick O'Farrell's group have cloned the *engrailed* gene, by chromosome-walking from a nearby tRNA gene, and shown that *engrailed* mutations define a 60 kb region. Again, this region gives rise to a number of discrete transcripts in a pattern that changes during development.

The cloning of the *Achaete/Scute* gene complex, reported by Juan Modellell (Madrid University), is of interest not only for its own sake, but also for the way in which it was achieved. Distributed about the surface of a fruit fly are innervated sensory organs called bristles or chaetes, of which there are two types: microchaetes, small and abundant; and macrochaetes, larger, fewer in number and in defined positions. *Achaete* mutations cause bald patches where there are no microchaetes, whereas *Scute* mutations lead to the loss of specific macrochaetes. *Achaete/Scute* double mutants have no chaetes whatsoever. The key to the cloning of the *Achaete/Scute* complex was a transposable element called 'gypsy', a copy of which appears to have inserted itself near the affected genes in all mutations suppressible by the recessive *suppressor of hairy-wing* mutation⁵.

Such a *gypsy* element in a suppressible *Scute/Achaete* mutation provided the starting place for a chromosome-walk that has now covered over 100 kb of this region. Five major transcripts have been mapped in the cloned region so far. It turns out there is a rough correspondence between the time of expression and position of the transcript on the map, with those seen earliest in development being encoded by sequences towards the end of the chromosome wherein map the strongest *Scute/Achaete* mutations. An interesting observation is that in mutants of a gene called *extra-macrochaetes* (phenotype self-explanatory), one of the major transcripts is observed at higher levels than normal, and the time of expression of another is prolonged.

What does it all mean?

The kinds of questions that molecular biologists will now have to confront may become clear from a consideration of the central role of homoeotic genes in development. A major problem is that of how each segment comes to express the correct gene combination. Garcia-Bellido⁶ has suggested that activator genes function early in development to regulate the expression of homoeotic genes, and an elegant series of

experiments carried out by Struhl have suggested that *extra-sex combs* is one such gene, regulating the expression of the *bithorax* complex^{7,8}. However, one is still left with the problem of how the overall body patterns are set up in the first place. Classical models have invoked gradients of 'morphogens', but so far no suitable candidate for such a role has come forward. Perhaps the application of the techniques described above to genes such as *Bicaudal*, mutations which produce embryos with two posterior poles, or *Dorsal*, which seems to regulate the setting-up of the dorsal-ventral polarity, will provide the answers to this problem.

Another problem is that of what homoeotic genes really do. They clearly make the decision as to what structures are made in each segment, but how do they do this? One might imagine they work by regulating a set of segment-specific 'realizator' genes⁶, but we have little idea how the complex patterns found within individual segments are formed. The overall pattern may be set up by spacial signals initiated at the boundaries between segments, and presumably cell-cell interactions play a crucial part: perhaps homoeotic genes help to determine the interactions between cell types.

In this regard, it may be worth noting that although the fates of epidermal cell lineages are determined early in development, other cell types are more plastic. Lawrence and Dan Brower have shown by transplanting genetically marked mesodermal cells that the structures they give rise to in the adult are determined not by their site of origin but by their site of transplantation⁹. Perhaps the rigid cuticular structures establish the basic plan, and more mobile cell types then simply match themselves to that plan.

A final note — is this of any relevance to organisms other than *Drosophila*? As Lawrence pointed out in answer to this very question, fundamental processes in biology tend to be general, despite the outward claims of different species to individuality. Those developmental biologists not fortunate enough to work with fruit flies may find all of this a bit depressing. □

*The British Society for Cell Biology Symposium on 'Genes controlling development' was held at Liverpool on 28 March 1983.

†The Nature conference to celebrate '30 years of DNA' was held at Churchill College, Cambridge on 13-14 April 1983 (see *Nature* 302, 651; 1983).

Geoffrey North is an assistant editor of *Nature*.

1. Garcia-Bellido, A., Lawrence, P.A. & Morata, G. *Sci. Am.* 241, 90 (1979).
2. Lewis, E.B. *Nature* 279, 565 (1978).
3. Struhl, G. *Proc. natn. Acad. Sci. U.S.A.* 79, 7380 (1982).
4. Mafen, E., Devine, N., Garber, R.L. & Gehring, W.J. *EMBO J.* 2, 617 (1983).
5. Modellell, J., Bender, W. & Meselson, M. *Proc. natn. Acad. Sci. U.S.A.* 80, 1678 (1982).
6. Garcia-Bellido, A. *CIBA Symp.* 29, 161 (1975).
7. Struhl, G. *Nature* 293, 36 (1981).
8. Struhl, G. & Brower, D.L. *Cell* 31, 285 (1982).
9. Lawrence, P.A. & Brower, D.L. *Nature* 295, 55 (1982).

A field guide to the binary stars

Virginia Trimble

Astronomy Program, University of Maryland, College Park, Maryland 20742, USA and
Department of Physics, University of California, Irvine, California 92717, USA

For most of the history of binary star astronomy, systems have been classified largely on the basis of how they were discovered and qualitative appearance of their spectra and light curves. Our understanding of single and double star evolution has now progressed to the point where most of the classes previously identified, and some new ones, can be arranged into evolutionary sequences, depending primarily on the initial masses and separation of the component stars.

OF the points of light in the sky we call stars, well over half in fact consist of two (or more) luminous bodies in gravitationally bound orbits around each other^{1,2}. These are the binary stars. About half of them, in turn, are close binaries—systems in which the stars are too close together to complete their normal evolution without being modified by each others' presence³.

For single stars, this normal stellar evolution can be crudely divided into several phases: (1) pre-main-sequence contraction from interstellar gas, (2) main sequence core hydrogen burning, ended by (3) exhaustion of hydrogen in the core, which contracts, initiating (4) shell hydrogen burning, during which the star appears as a red giant (or, for massive stars blue or yellow supergiant), followed by (5) core helium ignition and burning as horizontal branch or clump star or yellow-to-red supergiant, (6) shell helium burning as an asymptotic giant or red supergiant, and (7) rapid loss of outer layers (leaving a white dwarf) or rapid additional nuclear reactions ending in a supernova explosion and neutron star formation.

Each of these phases seems also to occur in binary systems and can be modified gently by the gravitational field or stellar wind of a companion or more drastically by gas flow to or from the companion through the inner lagrangian point (L_1 in Fig. 1) between them. This latter is called Roche lobe overflow. A star expands during phases (2), (4), and (6); and Roche lobe overflow beginning during them has historically³ been called Case A, B, and C of mass transfer respectively.

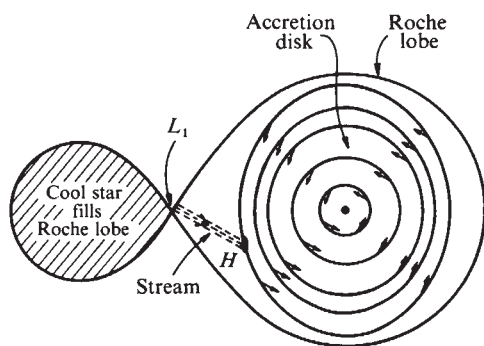


Fig. 1 Generalized, all-purpose close binary in process of mass transfer. Donor and receiver can each be in a variety of evolutionary states. Crucial aspects are the existence (in a coordinate frame that rotates with the system) of an equipotential surface that envelopes both stars, called the Roche lobe, and of the critical point L_1 on that lobe, between the stars, from which material will flow in a stream towards the receiver, carrying considerable angular momentum with it. Thus, at least during rapid transfer, a disk forms, and, especially in the cataclysmic variables, will have a hotspot (H) where the stream hits. Only slightly outside the Roche lobe occurs an equipotential that (again in the rotating frame) is open to the outside world, permitting ready loss of mass and angular momentum from the system as a whole.

A star of mass M evolves on a time scale roughly proportional to M^{-2} (main sequence lifetime $= (M/M_\odot)^{-2} \times 10^{10}$ yr, and so on). Thus, the initially more massive member of a pair, which we shall here and forever after call the primary, is always the first to reach the expansion phases and overflow its Roche lobe, if this is ever to happen at all. Thus we expect evolution of binary systems to occur in the following stages: (1) birth and pre-main-sequence contraction, (2) main sequence hydrogen burning in both stars, usually separate, but sometimes already in contact with their Roche lobes, (3) expansion of the primary, with first stage of mass transfer onto the secondary, and perhaps loss of both mass and angular momentum from the system as a whole, (4) completion of the evolution of the primary, (5) expansion of the secondary, with second stage of mass transfer back to primary and further losses from system, and (6) completion of evolution of the secondary. The sections below divide up the binary systems we have observed and modelled into these six classes, and Table 1 outlines the scheme.

We can also classify binary systems by the kind of observation that reveals the presence of the companion—changes in brightness, as one star eclipses, lights up, or distorts the shape of the

Table 1 Outline of the evolutionary scheme

Phase	What happens	Likely example
1	M_1 and M_2 contract	BM Ori
2	Both stars on main sequence	α Cen
	detached	W UMa
3	M_1 evolves and expands	RS Can Ven
	pre-contact	β Lyrae
	rapid transfer	Algol
	slow transfer	
4	M_1 completes evolution	KS Per
	helium stars	UU Sge
	binary planetary nebula nuclei	V471 Tauri
	white dwarf + main sequence	
5	M_2 evolves and expands	Mira
	pre-contact (M_1 = white dwarf)	
	symbiotic star	
	pre-contact (M_1 = neutron star)	Cen X-3
	massive X-ray binary	
	contact (M_1 = white dwarf)	Nova Cygni 1975
	cataclysmic variables	
	contact (M_1 = neutron star)	HZ Herc
	low-mass X-ray binary	
6	M_2 completes evolution	AM CVn
	two white dwarfs	Tycho's and Kepler's
	Type I supernova	Supernovae
	white dwarf + neutron star	long-period binary
	neutron star + neutron star	pulsar
		binary pulsar
		1913 + 16

other (eclipsing binaries), changes in radial velocity caused by orbital motion (spectroscopic binaries), discordant combinations of spectral lines coming from the two stars (spectrum binaries), and motion in the plane of the sky (astrometric binaries, visual binaries, and common proper motion pairs, for one dot of light wiggling, two dots orbiting, and two dots moving together in very wide orbit). This division is clearly less fundamental than the evolutionary one, but it is worth noting that we can know we have a binary system only if one of these effects is detected.

The literature of binary stars is, inevitably, enormous. In an effort to render this feast less indigestible, the present discussion has been largely tied to a secondary literature of review articles, monographs, and conference proceedings (listed as references 1, 3–24 and referred back to by many of the subsequent items). This serves two purposes: first, most of the points made will have been thought about by at least two people besides the present author, and second, the reader pursuing one of the items will generally find it nestled in context among other papers addressing related issues.

Birth and pre-main-sequence contraction (phase 1)

Star formation in general is rather poorly understood²⁵ and that of binaries doubly so²⁶, although we have some observational and theoretical evidence that there are two or more mechanisms at work, producing systems with different separations, mass ratios, and so on^{26,27}. We catch only a few systems during pre-main-sequence contraction, partly because the phase occupies only about 10^{-3} of the lifetime of a typical star, and partly because it is characterized by the presence of obscuring gas and dust clouds and by activity on the stellar surface that changes the brightness of the stars erratically and messes up their spectral lines so as to make detection of companions difficult.

T Tauri itself, the prototype pre-main-sequence variable, has a faint IR companion²⁸ with orbit period in excess of 1,000 yr, and a couple other members of the class show visual companions, eclipses, or variable radial velocities^{29–31}, the statistics of the velocities being consistent with normal binary incidence²⁹. Among more massive systems, a few like BM Ori³², seem to combine a pre-main-sequence secondary with a primary already burning hydrogen. Gaposchkin³⁰ lists some additional systems probably of this type, as well as systems containing UV Ceti variables, faint, flaring, low-mass stars, generally thought to be newly-arrived on the main sequence. The presence of a companion can, however, enhance flare activity³³, so that these need not all be very young. The BY Draconis variables, binaries with periods of a few days and extensive dark spots on the component stars, probably also have their activity enhanced this way³⁴.

The preceding remarks apply to Population I stars, those of roughly solar composition and ages $\leq 8,000$ Myr, found in the disk of our Galaxy. We just barely survey the brightest individual Population I stars in other nearby galaxies (Andromeda and the Magellanic Clouds), and these include eclipsing binaries in the same types and numbers as found in our own galaxy^{20,35}. The situation for the old, metal-poor stars of the halo (Population II) is quite different. A few isolated halo stars (identified as such by their composition and/or velocity) are spectroscopic binaries^{36,37}, but within the main concentration of halo stars, the globular clusters, we see no evidence at all for main sequence or giant binaries (although there are several X-ray binaries and, probably, cataclysmic variables, belonging to phase 5 below). This is not easy to explain, except by saying that main sequence systems never formed and the evolved systems must have arisen other than through the course of events just described³⁸. Our one other Population II-like sample comes from the Draco dwarf spheroidal galaxy, which apparently has no eclipsing binaries among its brightest stars^{20,35}. This situation is not only puzzling but annoying, as models for the dynamical evolution of globular clusters typically

make considerable use of energy transfer between binary orbits and the total cluster potential³⁹.

The main sequence (phase 2)

Stars spend 90% of their nuclear-burning lives on the main sequence. Catalogues of binary systems are rather less dominated than one might have guessed by systems containing two main sequence stars only because other, later phases are brighter and more spectacular. But we do see large numbers of main sequence pairs, with separations from nearly as large as half way to the next star⁴⁰ to as small as the sum of the stars' own sizes⁴¹.

Well-separated pairs are our only fundamental source of measured stellar masses and one of the most important sources of measured radii¹⁵. Thus binaries provide the foundation on which we have erected our entire understanding of stellar evolution. The measured masses range from 7% of the mass of our Sun for Ross 614 to a well-determined $32 \pm M_{\odot}$ for LY Aur¹⁵ and, perhaps, as much as $\geq 50 M_{\odot}$ for Plaskett's star²⁰ and V505 Mon⁴². A few astrometric companions may have still smaller masses and are, perhaps, not true hydrogen-burning stars⁴⁰; and some single stars may be even more massive than Plaskett's star⁴³. In addition to masses and sizes, we learn something about stars' interior density distributions from the way tidal interactions between binary companions change their mutual orbits. The observed distributions are in reasonable accord with the best model predictions⁴⁴.

Within detached systems, many of the things that can happen to single main sequence stars proceed unimpeded. For instance, the classes of mild variables dignified by the names β Cephei stars (or β CMa stars), α CVn stars, δ Scuti stars, and dwarf Cepheids are neither over nor under represented in binaries³², although some of the binary Delta Scuti variables probably have the precise frequencies of their pulsation modes modified by their companions' tidal influence⁴⁵. The interaction between duplicity and surface chemical abundance peculiarities is more complex: the metallic-line Am stars are essentially always short-period spectroscopic binaries, while the Ap stars (with strong lines of certain rare earths) almost never are^{46–48}. The Ap stars, but not the Am's, show strong surface magnetic fields⁴⁹. Both are slow rotators and can be modelled rather well by diffusion processes in their resultingly stable atmospheres^{50,51}. We might then blame the differences between the two classes on the effects of interactions among rotation, magnetic field, and companion⁴⁸.

This is perhaps the logical point to discuss systems so wide that both stars can complete their evolution without much interaction. Some of these are very informative, for instance the eclipsing class (with VV Ceph and ζ Aur as prototypes), in which the primary has developed an enormously extended, thin red-giant atmosphere, whose structure we can probe as the main-sequence secondary disappears gradually behind it⁵². Popper¹⁵ lists some completely detached red giant pairs; and two of the nearest stars, Sirius and Procyon, have white dwarf companions at such large distances that they must be the products of primaries that completed their evolution essentially unaffected by the secondaries we see now. In the 12 double-white dwarf common proper motion pairs⁴¹ both stars have died undisturbed. From an evolutionary point of view, these are really single stars, and will not concern us further here.

A little less than 0.1% of all stars⁵³ form in binaries of such small separation that both stars come in contact with their Roche lobes while on the main sequence. These are the W Ursa Majoris stars of spectral types *F*, *G*, and *K*, and their more massive analogues, the SV Cen stars. The latter are much the rarer, a reasonably complete list⁵⁴ including not many more than 12 well-studied systems. Contact binary specialists are currently disputing three items: how such systems form; how the component stars exchange mass, energy, and angular momentum to remain in contact; and how they eventually end up. An observed gap in separations between the W UMa's and the closest detached systems suggests that they do not quite

form in contact, but get there shortly after birth, perhaps losing angular momentum by magnetic braking⁵⁵. Energy exchange certainly occurs, as most pairs are more similar in luminosity than their mass ratios (as small as 0.08)⁵⁶ would permit for two single stars. The massive contact systems studied by Popper (personal communication) share this peculiarity. In SV Cen, the less massive star is actually the more luminous⁵⁷. Two models currently compete, one in which rather unstable transfer takes the systems repeatedly into and out of contact (on time scales too long for us to probe except statistically), and one in which contact is maintained steadily¹⁸. The former should lead to the stars' eventually coalescing⁵⁸, perhaps into a rapidly rotating giant like FK Comae⁵⁹, the latter possibly to their evolving into a cataclysmic binary (phase 5, below) configuration⁶⁰.

Evolution of the primary and first phase of mass transfer (phase 3)

As the primary begins to evolve and expand, effects attributable to the companion sometimes appear well before contact with the Roche lobe. Heating of the cooler component by the hotter one can drive some mass transfer⁶¹, and winds and other surface activity may be greatly enhanced. There are, for instance, the RS CVn variables, a numerous class of slightly evolved, solar-type stars, still well separated but, nevertheless, marked by conspicuous emission lines, and, quite unexpectedly, radio and X-ray emission of sorts not seen in most single stars of the same spectral types⁶². Most other radio-emitting stars are also binaries, the emission coming from enhanced winds, colliding winds, and the like⁶³, the chief exception being radio stars such as P Cygni which are known from their spectral line profiles to have strong winds anyway⁶⁴. X-ray stars, on the other hand (apart from the strong emitters discussed below), are typically single objects with extensive coronae⁶⁵. There is some disagreement on how much of the activity to blame on the presence of the companion and how much on rapid rotation and deep convection zones (D. M. Popper, personal communication).

As the primary reaches its Roche lobe, mass transfer begins on a Kelvin-Helmholtz time scale (because the lobe shrinks as the primary mass decreases), until the mass ratio is reversed, after which transfer continues on a nuclear time scale³.

We catch a few stars, such as ϵ Aurigae⁶⁶ and β Lyrae⁶⁷ during the short-lived phase of rapid transfer, and many more, the W Serpentis stars⁶⁸ just after the mass ratio is reversed. The brightest stars in other galaxies (the Hubble-Sandage variables) and our local example, η Carinae, are perhaps very massive binaries also experiencing rapid transfer⁶⁹. During the most rapid transfer, the secondary (like that of β Lyrae) is often completely concealed by a disk of accreting material. In fact, the secondaries are likely to have trouble accommodating both the mass⁷⁰ and the angular momentum^{71,72} as fast as they arrive, and considerable amounts of both can be lost to the system. We know that this happens: several systems show spectral evidence for shells surrounding both stars, gas streams away from them, and the like^{73,74}; and the masses and separations of pairs that have completed rapid transfer show that mass and angular momentum cannot generally have been conserved within the systems⁷⁵⁻⁷⁸.

The most severe losses occur when the expanding primary surrounds both stars in a common envelope⁷⁸. The stars, dragging on the envelope, spiral together and shove it off. Some close pairs among the X-ray binaries, cataclysmic variables, binary nuclei of planetary nebulae, and V471 Tauri stars must be products of such a process. Webbink⁷⁹ has attempted to trace some of the intermediate phases and examples of each. The sequence, as the envelope gradually takes over and then leaves, runs: recurrent nova, symbiotic star of the CI Cyg type⁸⁰, BQ[] star (B type star with spectrum dominated by forbidden emission lines)⁸¹, F or G emission line supergiant (like ρ Cass), ultra-long period contact binary (such as W Cru), rapidly rotating bright giant with double core (such as FK Comae, although

it and its relatives have alternatively been explained as the product of merged W UMa systems⁵⁹), star with strong emission and absorption lines showing rapid mass loss (such as P Cygni), and planetary nebula with double nucleus (such as NGC6164/65 and HD148937). Webbink does not mean to imply that all members of these observed classes are to be explained in this way; many recurrent novae and symbiotic stars, for instance, clearly belong in phase 5.

As mass transfer and loss slow down to the nuclear time scale, we reach the Algol binaries, in which the by now lower mass primary continues to drizzle material onto the more massive secondary sufficiently slowly that there is no conspicuous accretion disk⁸². The class includes Algol itself and many other well-studied stars, and explaining the seeming paradox of the lower mass but more highly evolved primary was one of the early triumphs of binary star theory³. Two puzzles remained. One, that of the R CMa and undersized subgiant stars (in which the primary either had remarkably low mass or seemed to have pulled away from its Roche lobe while maintaining an extended envelope), has now largely been eliminated by more accurate observations of the systems concerned⁸³. The other is still with us. Naftilan⁸⁴ among others found spectroscopic evidence that the evolved star had a lower surface abundance of heavy elements than its companion in many Algols. The implication is that the elements other than hydrogen and helium are concentrated towards stellar surfaces, and stripping has revealed the low-metal zone⁸⁵. This, if true, has enormous implications for models of chemical evolution of the Galaxy, for the solar neutrino problem, and so on⁸⁷. Recent work has not much clarified the situation: giant components have been reported with metal abundances that are lower than⁸⁷, the same as⁸⁸, and higher than⁸⁹ those of the unevolved stars. I hope this is all noise, not signal.

At least some of the Be and shell stars (prototype γ Cas) belong here^{5,21}. These show both rapid rotation and emission lines indicative of hot gas around them. A large fraction are undoubtedly in binary systems, and their shells attributable to material being accreted from a more highly evolved companion⁹⁰. The Be stars with X-ray emitting companions, such as X Per, must, on the other hand, be losing material to the neutron stars⁹¹ and belong in phase 5.

Demise of the primary (phase 4)

As the primary continues its evolution, helium ignition or loss of the entire hydrogen-rich envelope will reduce its size until it pulls away from its Roche lobe and mass transfer ceases (possibly to be briefly renewed during shell helium burning). The system becomes much less conspicuous, eventually simply appearing as a fairly normal main sequence star, whose white dwarf or neutron star companion may or may not be detectable. A sufficiently asymmetric supernova explosion of a massive primary can unbind its system, sending both neutron star and OB secondary off at high speed, the former as a pulsar, the latter as a classical OB runaway star⁹². But explosions that yield recoil without disruption must also occur, since about half of known OB runaways have low-mass companions still attached⁹³. About half of these companions should be white dwarfs or neutron stars, and half small main sequence stars⁹⁴, runaway velocities in the latter case being attributable to ejection from a cluster.

Systems that have reached this stage may reveal their wild past in several ways. First, the secondary will be more massive than a single star of the same age could be. The puzzling blue stragglers, members of galactic and globular clusters and dwarf spheroidal galaxies that are too massive (too blue and bright) to have spent the whole age of their parent star systems in their present state may be mass transfer products^{95,96}, although mixing between core and envelope in a single star can produce the same effect⁹⁷.

Second, either star or both can show composition anomalies as a result of the primary having been stripped down to where hydrogen or helium burning have gone on. Highly evolved, but

pre-white-dwarf, primaries can appear as helium stars (meaning, at least, $\text{He}/\text{H} > 1$) like ν Sgr and KS Per^{98,99}, some of these also showing excess nitrogen from CNO cycle hydrogen burning¹⁰⁰, or as Wolf-Rayet binaries²⁴, in which a compact, but very bright core with lots of He and N or C but no discernible hydrogen orbits a fairly normal OB main sequence star^{101,102}. Evolution and stripping of the secondary can eventually produce a second Wolf-Rayet binary stage with white dwarf companion, comprising 20–50% of the observed systems^{103–106}. These really belong in phase 6. And there are also single Wolf-Rayet stars, stripped down to the same levels by violent winds¹⁰⁷, but a binary companion clearly helps³. The barium stars may belong in here too. Barium is made by slow addition of neutrons to iron in parts of stars where hydrogen and helium burning can interact. Thus an excess of it means that a star's surface has been adulterated by its own or a companion's core nuclear burning products. Self-induced barium excesses appear among (typically single) carbon rich stars¹⁰⁸. But the classical giant (normal carbon) barium stars are apparently all binaries, with rather low mass secondaries¹⁰⁹. The obvious interpretation, that the pollution came from the companion, which is now a white dwarf, is not entirely supported by observations, and these systems are still rather a mystery¹¹⁰.

Third, we may still see portions of an envelope shed by the primary or by common envelope processes, as in the planetary nebula with eclipsing, spectroscopic, or visual binary nuclei like UU Sge, FG Sge, NGC 246, 1514, and 6543 (refs 111–113), and in the class of binary protoplanetary nebulae (prototypes HM Sge and V1016 Cyg), which may or may not really be either binaries or planetary nebula precursors^{114–115}. The supernova remnant G109.1–1.0, whose central object is an X-ray binary, is another example of a recent demise (to neutron star rather than white dwarf) with remaining gas¹¹⁶.

Finally, the system may just sit there and stare at you like the V471 Tauri stars, a handful of systems known to have white dwarf primaries, low mass, main-sequence secondaries, and separations small enough that they must be common envelope products^{78,117,163}. An extreme example, AA Dor, has a total remaining mass of less than $0.3 M_{\odot}$ (ref. 118) and must have lost mass and angular momentum with astounding efficiency, perhaps by magnetic braking and winds as well as a common envelope (P. P. Eggleton, personal communication). None of these has the secondary main sequence mass significantly larger than the white dwarf, as would result from conservative mass transfer in a system with initially almost equal components.

Evolution of the secondary (phase 5)

In due course, the secondary in turn departs the main sequence, resulting in mass transfer back to the primary. Occasionally this may happen before the primary has completed its evolution, resulting in a pair of helium stars, which can either end up as a close white dwarf pair (like others in phase 6) or, with large enough initial masses, give rise to a nuclear-detonation Type I supernova¹¹⁹.

Normally, the primary is compact before back transfer begins, just because of the steep dependence of stellar lifetime on mass¹²⁰. It thus has a deep gravitational potential well, and the secondary can contribute enough accretion just from its wind to produce conspicuous effects before it ever reaches its Roche lobe. Systems such as Mira (pulsating red giant + white dwarf in a wide orbit) are of this type, as are some of the symbiotic stars (such as V1329 Oph)¹²². This class of erratic variables²³ is defined by the simultaneous presence of spectral features coming from regions with two very different temperatures. Some may be single stars with extended coronae, but most (and all 'real' symbiotics, in current usage) are binaries with significant accretion, normally onto a white dwarf, from wind or Roche lobe overflow. Drilling and Schonberger¹²² suggest that the helium-rich binaries KS Per and ν Sgr also involve slow transfer back onto a compact primary. Finally, where a massive O or B secondary's wind impinges on a neutron star, we see a binary X-ray source.

Once the secondary reaches its Roche lobe, transfer speeds up. For a massive secondary and neutron star primary, it becomes so rapid that ambient gas is optically thick to X rays, and the source turns off. For a lower-mass secondary, on the other hand, transfer now for the first time becomes sufficient to produce observable X rays, so that we see two separate, high and low mass, classes of X-ray binaries. These must have been produced, respectively, by first-stage transfer that lost very little and much mass from the system as a whole^{123,124}. The mass gap in between represents stars where the wind transfer rate is too small and the Roche lobe overflow rate too large to make detectable X rays. The neutron stars in most of these systems have rotation periods of seconds to minutes (minimum 0.069 for A0538–66)¹²⁵, surprisingly slow given that accretion is adding angular momentum and the radiation mechanism not draining it away as in true pulsars. Spin-down periods must alternate with accretion episodes in some way¹²⁶ or we could not catch as many systems as we do rotating slowly and spinning up¹²⁷. Interesting special cases of the X-ray binary phase include 4U1626–67 (ref. 128), whose donor secondary is apparently a white dwarf; SS433 (ref. 129), in which some of the energy goes into expelling two oppositely directed gas jets at about $0.26c$; and Cyg X-1, whose primary is almost certainly a black hole¹⁰. Continuing accretion onto a neutron star in a close binary is an obvious and straightforward way to form a black hole, but the Cyg X-1 primary is too massive to have grown by this mechanism during the lifetime of the OB secondary, and must have resulted directly from the collapse of a star unable to eject all its envelope in a supernova ($M \geq 15M_{\odot}$ according to Hillebrandt¹³⁰).

Finally, Roche lobe overflow onto a white dwarf primary results in an enormous range of phenomena collectively dignified by the name cataclysmic binaries¹³¹. The main subtypes are the classical novae, the dwarf novae (or SS Cygni stars, subtypes named for Z Cam, with luminosity plateau on the declining branch of the light curve, and U Gem without it), the nova-like variables (UX UMa stars), the polars (or AM Herc stars), the recurrent novae, and the symbiotic (or Z And) stars. Many of them are X-ray sources¹³², though never so bright or hard as the accreting neutron stars. The division among types is not absolutely rigid, the known recurrent novae T CrB and RS Oph looking, in between outbursts, like symbiotic stars¹³³.

Table 2 summarizes observed properties of the several types and what we think we know about the nature of the component stars and the mode of mass transfer. The basic energy source is gravitational (accretion) for the outbursts of the dwarf novae¹³⁴ and nuclear (degenerate ignition of hydrogen) for the novae and recurrent novae¹³⁵. This, plus the size of the donor star, the rate of mass transfer, and intensity of magnetic field of white dwarf seem to divide up the various types (refs 136–142, and B. Paczyński, personal communication). We do not identify all possible combinations, including the lowest possible transfer rates (presumably because nothing interesting happens), and the highest ones, at which the hydrogen burns steadily and the systems presumably are indistinguishable from hydrogen-shell-burning giants. These high transfer rates may conceal the systems in which the white dwarf is still the less massive star. These should otherwise be fairly numerous (B. Paczyński, personal communication). Considerable work remains to be done in sorting out the physics of the several classes, aberrant members, and, especially, the precise progenitors, evolutionary history, and cause of mass transfer for each^{4,6,8,14,16,23,143}.

Several kinds of mass ejection occur in cataclysmic variables: 'true' novae and recurrent novae blow off the partially-burned hydrogen of each explosion in a detectable, expanding remnant¹⁴⁴. The jets ejected by R Aqr¹⁴⁵, the nucleus of planetary nebula A30¹⁴⁶ and some of the Herbig-Haro objects¹⁴⁷ may or may not have anything to do with this phase. They are all rather similar looking, but the first is almost certainly a binary, and the last almost certainly not.

Table 2 The cataclysmic variables classified in the conventional way

Type	M_v (quiescent)	Outburst amplitude (mag)	Recurrence time	Nature of instability	Stars	Mass transfer
Novae	+5	9–14	10^{4-5} yr	Degenerate ignition of hydrogen	Main sequence + white dwarf	Lobe overflow $\sim 10^{-9} M_\odot \text{ yr}^{-1}$
Dwarf novae	+10	2–6	10 days–30 yr	Change in $\dot{M}$ and/or disk structure	Main sequence + white dwarf	Lobe overflow $\sim 10^{-10} M_\odot \text{ yr}^{-1}$
Nova-like	+5–10	Irregular	—	Change in $\dot{M}$ and/or disk structure	Main sequence + white dwarf	Lobe overflow $\sim 10^{-9-10} M_\odot \text{ yr}^{-1}$
Polar	+5–10	Irregular	—	Change in $\dot{M}$ and/or accretion pattern	Main sequence + white dwarf strong B	Lobe overflow $\sim 10^{-9-10} M_\odot \text{ yr}^{-1}$
Recurrent novae	+2	7–9	10–100 yr	Degenerate ignition of hydrogen	Red giant + white dwarf	Lobe overflow $\sim 10^{-6} M_\odot \text{ yr}^{-1}$
Symbiotic stars	+4–0	Irregular	—	Changes in pattern of quasi-spherical accretion	Red giant + white dwarf	Wind $\sim 10^{-7} M_\odot \text{ yr}^{-1}$

Cataclysmic variables and X-ray binaries, unlike previous stages, are well represented among Population II stars in globular clusters. The strong cluster X-ray sources appear to be standard accreting neutron stars¹⁹, and their number (about eight at last count¹⁴⁸) constitutes far more than the clusters' fair share of the galactic supply¹⁴⁹. Optically-identified cataclysmic variables are perhaps similarly overabundant³⁸ and still more cataclysmic variables are the likeliest explanation of a new class of weak ($L = 10^{32-34} \text{ erg s}^{-1}$) cluster X-ray source which currently has about 12 members^{148,150}.

As during the first transfer phase, donors may now envelop their companions, resulting in extensive mass and angular momentum loss from the system. Some of the closest 'dead' pairs described below require something of the sort within their history^{151,152}.

For the cataclysmic variables especially, the death process can be gradual and fairly prolonged, resulting in pairs such as HZ 22^{153,154}, where a hot subdwarf (pre-white dwarf in effect) orbits a cooler degenerate star with very little interaction, and AM CVn¹³⁷ whose components are both low mass degenerates but, owing to their very close spacing (orbital period 18 min), interact vigorously enough to put the system on many lists of nova-like variables.

A spectacular termination is also possible: prolonged accretion of burned hydrogen can drive the white dwarf to one of several kinds of violent burning of its entire helium, carbon, and oxygen supply. This is our current 'best-buy' model for (many or possibly all, though progenitors are rather rare¹⁵⁵) Type I supernovae, the sort occurring among relatively old stars¹⁵⁶.

Death of the secondary (phase 6)

The secondary has much the same options available to it as the primary—planetary nebula plus white dwarf (UU Sge may be such a second-time-around planetary nebula with binary nucleus⁷⁹), neutron star, or black hole. The first of these leads typically to a pair of white dwarfs, with separation depending on the amount of angular momentum lost in previous stages. These are inevitably inconspicuous, but we see a few pairs, both close (AM CVn) and wide (for example, G107–69/70)¹⁵⁷. Many others may be hiding among seemingly-single white dwarfs and could be identified by careful searches for colour anomalies and variable radial velocities. Loss of angular momentum by gravitational radiation and, perhaps, winds, inevitably produces coalescence of such a system, though not necessarily within the age of the Universe. The result is likely to look like a Type I supernova (B. Paczynski, personal communication).

Less often, the secondary white dwarf may find itself orbiting a primary neutron star. The two longer-period binary pulsars may be like this^{158,159}. We know the neutron stars must have formed first, because their combination of rapid rotation and weak magnetic field requires them to have been spun up by

accretion from a companion¹⁶⁰. The one visible star within the error box of the 24-h orbit period example is, however, fainter than this hypothesis predicts. Loss of sufficient material from the secondary in a planetary nebula could, in principle, release the newly-spun-up neutron star. Arons¹⁶¹ proposes this origin for the new 1.56-ms pulsar¹⁶².

Transition of the secondary to a neutron star may lead to a Type II supernova, with sufficient ejection to unbind the (now less massive) primary, yielding an old and a young runaway neutron star, with suitable velocities to explain the pulsar speeds we see. The system may also remain bound if the secondary has already lost most of its hydrogen and helium layers in a common envelope phase. The supernova event could then be quite faint (and/or look like a Type I, owing to the lack of hydrogen), and the product resemble the best-known binary pulsar, 1913+16, whose companion is probably another neutron star^{158,159}. Which is the older is puzzling, since spin-up of the one we see apparently required accretion from the companion¹⁶⁰, which we then might also expect to see as a pulsar.

Finally, if the secondary collapses to a black hole, we do not expect to see anything at all from the black hole + black hole or black hole + neutron star pair, short of a final burst of gravitational radiation as they spiral together, or a compressed-dentifrice tube squirt of heavy elements from the disappearing neutron star.

Most of this material was collected in the libraries of the Institute of Astronomy, Cambridge, whose director, Professor Donald Lynden-Bell, I thank for hospitality. I first came to think seriously about binary stars under the influence of Drs Daniel M. Popper and Bohdan Paczynski. This paper is dedicated to the memory of John A. J. Whelan.

1. Batten, A. H. *Binary and Multiple Systems of Stars* (Pergamon, Oxford, 1973).
2. Michell, J. *Phil. Trans. R. Soc.* 234ff (1767).
3. Paczynski, B. *A. Rev. Astr. Astrophys.* 9, 183 (1971).
4. Robinson, E. L. *A. Rev. Astr. Astrophys.* 14, 119 (1976).
5. Slettebak, A. (ed.) *Be and Shell Stars*, IAU Symp. No. 70 (1976).
6. Eggleton, P. P., Mitton, S. & Whelan, J. A. J. (eds) *Structure and Evolution of Close Binaries*, IAU Symp. No. 73 (1976).
7. Thomas, H.-C. A. *Rev. Astr. Astrophys.* 15, 127 (1977).
8. Friedjung, M. (ed.) *Novae and Related Stars* (Reidel, Dordrecht, 1976).
9. Gallagher, J. S. & Starrfield, S. A. *Rev. Astr. Astrophys.* 16, 171 (1978).
10. Bahcall, J. N. A. *Rev. Astr. Astrophys.* 16, 241 (1978).
11. Sahade, J. & Wood, F. B. *Interacting Binary Stars* (Pergamon, Oxford, 1978).
12. Kopal, Z. *Dynamics of Close Binaries* (Reidel, Dordrecht, 1978).
13. Conti, P. S. & de Loore, C. W. H. (eds) *Mass Loss and Evolution of O Type Stars*, IAU Symp. No. 83 (1979).
14. Van Horn, H. & Wiedemann, V. (eds) *White Dwarfs and Variable Degenerate Stars*, IAU Colloq. No. 53 (1979).
15. Popper, D. M. A. *Rev. Astr. Astrophys.* 18, 115 (1980).
16. Plavec, M. J., Popper, D. M. & Ulrich, R. K. (eds) *Close Binary Stars: Observations and Interpretation*, IAU Symp. No. 88 (1980).
17. Shu, F. H. & Lubow, S. H. A. *Rev. Astr. Astrophys.* 19, 299 (1981).
18. Sugimoto, D., Lamb, D. Q. & Schramm, D. N. (eds) *Theory of Stellar Evolution*, IAU Symp. No. 93 (1981).
19. Holt, S. S. & McCray, R. A. *Rev. Astr. Astrophys.* 20, 233 (1982).
20. Herczeg, T. in *Landolt-Börnstein*, New Ser. Vol. 2, 381 (Springer, Berlin, 1982).
21. Jaschek, M. & Groth, H.-G. (eds) *Be Stars*, IAU Symp. No. 98 (1982).
22. Kopal, Z. & Rahe, J. (eds) *Binary and Multiple Systems of Stars as Tracers of Stellar Evolution*, IAU Colloq. No. 69 (1982).
23. Friedjung, M. & Viotti, R. (eds) *The Nature of Symbiotic Stars*, IAU Colloq. No. 70 (1982).

24. de Loore, C. W. H. & Willis, A. J. (eds) *Wolf Rayet Stars*, IAU Symp. No. 99 (1982).
25. De Jong, T. (ed) *Star Formation*, IAU Symp. No. 75 (1977).
26. Lucy, L. *IAU Symp.* No. 88, 7 (1980).
27. Trimble, V. & Cheung, C. *IAU Symp.* No. 70, 369 (1976).
28. Zuckerman, B. *Colloquium on the Infrared Companion of T Tauri*, Univ. of Maryland (1981).
29. Jones, B. F. & Herbig, G. *Astr. J.* **84**, 1872 (1979).
30. Gaposchkin, C. H. P. A. *Rev. Astr. Astrophys.* **16**, 1 (1978).
31. Bakos, G. A. & Tremko, J. *IAU Colloq.* No. 69, 67 (1982).
32. Sahade, J. & Wood, F. B. *Interacting Binary Stars*, 74 (Pergamon, Oxford, 1978).
33. Stauffer, J. *Astr. J.* **87**, 899 (1982).
34. Bopp, B. *Highlights Astr. S.* **8**, 447 (1980).
35. Herczeg, T. J. *IAU Colloq.* No. 69, 145 (1982).
36. Bolton, C. T. & Thompson, J. R. *IAU Symp.* No. 88, 575 (1980).
37. Mayer, M. & Turon, C. *Astr. Astrophys.* **110**, 241 (1982).
38. Trimble, V. *IAU Symp.* No. 85, 259 (1980).
39. Heggie, D. L. *IAU Symp.* No. 85, 401 (1980).
40. van de Kamp, P. *IAU Colloq.* No. 69, 67 (1982).
41. Shapley, H. *Harvard Observatory Monogr.* **7**, 249 (1948).
42. Stagni, R., Margoni, R. & Mammano, A. *Astrophys. Space Sci.* **88**, 115 (1982).
43. Maeder, A. *Astr. Astrophys.* **92**, 101 (1981).
44. Giménez, A. & García-Pelayo, J. *IAU Colloq.* No. 69, 37 (1982).
45. Antonello, E. *IAU Colloq.* No. 69, 33 (1982).
46. Abt, H. *Vistas Astr.* **8**, 75 (1965).
47. Preston, G. W. A. *Rev. Astr. Astrophys.* **12**, 257 (1974).
48. Kitamura, M. *IAU Symp.* No. 88, 95 (1980).
49. Borra, E. F., Landstreet, J. D. & Mestel, L. A. *Rev. Astr. Astrophys.* **20**, 191 (1982).
50. Michaud, G., Megassari, C. & Charland, Y. *Astr. Astrophys.* **103**, 264 (1981).
51. Vauclair, S. & Vauclair, G. A. *Rev. Astr. Astrophys.* **20**, 37 (1982).
52. Sahade, J. & Wood, F. B. *Interacting Binary Stars*, 120 (Pergamon, Oxford, 1978).
53. Budding, E. *IAU Colloq.* No. 69, 351 (1982).
54. Leung, K.-C. *IAU Symp.* No. 88, 257 (1980).
55. Rahunen, T. & Vilhu, O. *IAU Colloq.* No. 69, 289 (1982).
56. Mianowski, O. T. in *Photometric and Spectroscopic Binary Systems* (eds Carling, E. B. & Kopal, Z.) (Reidel, Dordrecht, 1981).
57. Sugimoto, D. & Miyagi, S. *IAU Symp.* No. 88, 133 (1980).
58. Webbink, R. *IAU Symp.* No. 88, 477 (1980).
59. Bopp, B. W. & Rucinski, S. M. *IAU Symp.* No. 93, 177 (1980).
60. Shu, F. H. *IAU Symp.* No. 88, 477 (1980).
61. Kondo, Y. & Modisette, J. L. *IAU Colloq.* No. 69, 317 (1982).
62. Popper, D. M. *IAU Symp.* No. 88, 387 (1980).
63. Gibson, D. M. *IAU Symp.* No. 88, 31 (1980).
64. Wendkar, H. J., Baars, J. W. M. & Altenhoff, W. J. *Nature phys. Sci.* **245**, 118 (1973).
65. Vaina, G. S. *Space Sci. Rev.* **30**, 151 (1981).
66. Plavec, M. J. *IAU Colloq.* No. 69, 159 (1982).
67. Wilson, R. E. *IAU Colloq.* No. 69, 261 (1982).
68. Plavec, M. J. *IAU Symp.* No. 88, 251 (1980).
69. Bath, G. T. *IAU Symp.* No. 88, 155 (1980).
70. Benson, R. S. thesis, Univ. California, Berkeley (1970).
71. Wilson, R. E. & Trigg, L. W. *IAU Symp.* No. 88, 263 (1980).
72. Packet, W. *Astr. Astrophys.* **102**, 19 (1981).
73. Polidan, R. S. & Peters, G. J. *IAU Symp.* No. 88, 293 (1980).
74. Sahade, J. & Wood, F. B. *Interacting Binary Stars*, 47, 73 (Pergamon, Oxford, 1978).
75. Smak, J. *IAU Symp.* No. 88, 575 (1980).
76. De Grève, J. P. & Vanbeveren, D. *IAU Symp.* No. 88, 103 (1980).
77. Giuricin, G., Mardirossian, F. & Mezzetti, M. *IAU Symp.* No. 88, 207 (1980).
78. Paczyński, B. *IAU Symp.* No. 73, 65 (1976).
79. Webbink, R. F. *Changing Trends in Variable Star Research*, IAU Colloq. No. 46, 102 (1980).
80. Bath, G. & Pringle, J. *Mon. Not. R. astr. Soc.* **201**, 345 (1982).
81. Ciatti, F., D'Odorico, S. & Mammano, A. *Astr. Astrophys.* **34**, 181 (1974).
82. Kaichuk, R. H. & Honeycutt, R. K. *Publ. astr. Soc. Pacif.* **94**, 532 (1982).
83. Hall, D. S. & Neff, S. G. *Acta Astr.* **29**, 641 (1979).
84. Naftilan, S. *Publ. astr. Soc. Pacif.* **87**, 321 (1975).
85. Bond, H. E. *Publ. astr. Soc. Pacif.* **84**, 839 (1972).
86. Trimble, V. *Rev. Mod. Phys.* **47**, 906 (1975).
87. Kondo, M. & Okazaki, A. *IAU Symp.* No. 88, 221 (1980).
88. Fekel, F., Lacy, C. H. & Tomkin, J. *IAU Symp.* No. 88, 53 (1980).
89. Olsen, E. H. *Astr. Astrophys.* **110**, 215 (1982).
90. Peters, G. J. *IAU Symp.* No. 88, 215 (1982).
91. Crampton, D. *IAU Symp.* No. 88, 313 (1980).
92. de Cuyper, J.-P. *IAU Colloq.* No. 69, 417 (1982).
93. Stone, R. C. *Astr. J.* **87**, 90 (1982).
94. Meurs, E. J. A., van den Heuvel, E. P. J. & Reusent, R. *Astr. Astrophys.* (submitted).
95. McCrea, W. H. *Mon. Not. R. astr. Soc.* **128**, 147 (1964).
96. Renzini, A., Mengel, J. G. & Sweigart, A. V. *Astr. Astrophys.* **56**, 369 (1977).
97. Wheeler, J. C., Saio, H. & Breger, M. *IAU Symp.* No. 93, 187 (1981).
98. Plavec, M. J. *IAU Symp.* No. 51, 216 (1973).
99. Bidelman, W. P. *IAU Symp.* No. 83, 395 (1979).
100. Bolton, C. T. *IAU Symp.* No. 83, 308 (1979).
101. Massey, P. *IAU Symp.* No. 99, 251 (1982).
102. Vanbeveren, D. & de Loore, C. *IAU Symp.* No. 88, 115 (1980).
103. Vanbeveren, D. *IAU Symp.* No. 88, 169 (1980).
104. Moffat, A. F. J. & Seggewiss, W. *IAU Symp.* No. 88, 181 (1980).
105. Cherepashchuk, A. M. *Astrophys. Space Sci.* **86**, 299 (1982).
106. Vanbeveren, D., Van Rensberger, W. & de Loore, C. *Astr. Astrophys.* **115**, 68 (1982).
107. Chiosi, C. *IAU Symp.* No. 99, 323 (1982).
108. Trimble, V. *Rev. Mod. Phys.* **47**, 902 (1975).
109. McClure, R. D. Preprint, *The Binary Nature of Barium Stars II. Velocities, Binary Frequencies, and Preliminary Orbits* (Dominion Astrophysical Observatory, 1982).
110. Dominy, J. F. & Lambert, D. L. Preprint, *Do All Barium Stars have a White Dwarf Companion?* (University of Texas, Austin, 1982).
111. Sahade, J. & Wood, F. B. *Interacting Binary Stars*, 118 (Pergamon, Oxford, 1978).
112. Acker, A. thesis, Publ. Obs. Astron. Strasbourg **4**, fasc. 1 (1976).
113. Mendez, R. H. *IAU Symp.* No. 88, 567 (1980).
114. Blair, W. D., Stencel, R. E., Shaviv, G. & Feibelman, W. A. *Astr. Astrophys.* **99**, 73 (1981).
115. Tarahora, O. G. & Yudin, B. F. *Soviet Astr. AJ Lett.* **8**, 46 (1982).
116. Gregory, P. C. *Pap. at XVIIIth General Assembly of International Astronomical Union*, Patra, Greece (Session of Commission 42, 1982).
117. Paczyński, B. & Dearborn, D. S. P. *Mon. Not. R. astr. Soc.* **190**, 395 (1980).
118. Hilditch, R. W., Hill, G. & Kilkenny, D. *IAU Colloq.* No. 69, 455 (1982).
119. De Grève, J. P. *Astrophys. Space Sci.* **84**, 447 (1982).
120. de Loore, C. *Space Sci. Rev.* **26**, 113 (1980).
121. Chacal, D., Grygar, J. & Hric, L. *IAU Symp.* No. 88, 547 (1980).
122. Drilling, J. S. & Schönberner, D. *Advances in UV Astronomy* (NASA CP-2238, 1982).
123. van den Heuvel, E. P. J. *IAU Symp.* No. 95, 379 (1981).
124. van den Heuvel, E. P. J. *IAU Symp.* No. 93, 155 (1981).
125. Skinner, G. K. *et al.* *IAU Circ.* No. 3671 (1982).
126. Ziolkowski, J. *IAU Symp.* No. 88, 335 (1980).
127. Lipunov, V. M. *Astrophys. Space Sci.* **82**, 343 (1982).
128. Crampton, D. *IAU Symp.* No. 88, 313 (1980).
129. *Vistas Astr.* **25**, 1981
130. Hillebrandt, W. *Ann. N.Y. Acad. Sci.* (in the press).
131. Payne-Gaposchkin, C. H. *Astr. J.* **82**, 665 (1977).
132. Cordova, F. A. & Mason, K. O. in *Stellar Driven Accreting X-Ray Sources* (eds Lewin, W. H. G. & van den Heuvel, E. P. J.) (Cambridge University Press, 1983).
133. Sahade, J. & Wood, F. B. *Interacting Binary Stars*, 117 (Pergamon, Oxford, 1978).
134. Bath, G. T. *Q. J. R. astr. Soc.* **19**, 442 (1978).
135. Lebedinsky, A. & Gürevich, L. E. *Dokl. Akad. Nauk SSSR* **55**, No. 9 (1947).
136. Mallama, A. & Trimble, V. *Q. J. R. astr. Soc.* **19**, 430 (1976).
137. Robinson, E. L. A. *Rev. Astr. Astrophys.* **14**, 119 (1976).
138. Warner, B. *IAU Symp.* No. 73, 85 (1976).
139. Paczyński, B. & Rudak, B. *Astr. Astrophys.* **82**, 349 (1980).
140. Nomoto, K., Narai, K. & Sugimoto, D. *Publ. astr. Soc. Jap.* **31**, 287 (1979).
141. Chiappetti, L., Tanzi, E. G. & Treves, A. *Space Sci. Rev.* **27**, 3 (1980).
142. Plavec, M. J. *IAU Colloq.* No. 70, 231 (1982).
143. Ritter, H. Cataclysmic binaries and related objects in the context of stellar evolution (Nanking Workshop on High Energy Astrophysics, in the press).
144. Barnard, E. *Astr. J.* **30**, 86 (1917).
145. Wallerstein, G. & Greenstein, J. L. *Publ. astr. Soc. Pacif.* **92**, 275 (1980).
146. Hazard, C., Terlevich, R., Morton, D. C., Sargent, W. L. W. & Ferland, G. *Nature* **285**, 463 (1980).
147. Herbig, G. & Jones, B. F. *Astr. J.* **86**, 1232 (1981).
148. Grindlay, J. *Colloq.*, Univ. of Maryland (1982); Hertz, P. & Grindlay, J. *Bull. Am. astr. Soc.* **18**, 888 (1982).
149. Katz, J. *Nature* **253**, 698 (1975).
150. Bode, M. F., Evans, A. & Bruch, A. *IAU Colloq.* No. 69, 475 (1982).
151. Paczyński, B. *Highlights Astr.* **5**, 27 (1979).
152. Paczyński, B. in *X-Ray Astronomy* (eds Baity, W. A. & Peterson, L. E.) (Pergamon, Oxford, 1979).
153. Greenstein, J. L. *Astr. Astrophys.* **23**, 1 (1973).
154. Schönberner, P. *Astr. Astrophys.* **70**, 451 (1978).
155. Trimble, V. *Observatory* **102**, 133 (1982).
156. Rees, M. J. & Stoneham, R. J. (eds) *Supernovae* (Reidel, Dordrecht, 1982).
157. van de Kamp *IAU Colloq.* No. 69, 81 (1982).
158. Taylor, J. H. *IAU Symp.* No. 95 (1981).
159. Blanford, R. D. & DeCampli, W. M. *IAU Symp.* No. 95, 371 (1981).
160. Taylor, J. H. *Ann. N.Y. Acad. Sci.* (in the press).
161. Arons, J. *Ann. N.Y. Acad. Sci.* (in the press).
162. Backer, D. C., Kulkarni, S. R., Heiles, C., Davies, M. M. & Goss, W. M. *Nature* **300**, 615 (1982).
163. Paczyński, B. *Acta Astr.* **30**, 113 (1980).

ARTICLES

Pb-Sr isotope variation in Indian Ocean basalts and mixing phenomena

Bernard Dupré & Claude J. Allègre

Laboratoire de Géochimie-Cosmochimie (LA 196), Institut de Physique du Globe et Département des Sciences de la Terre, Université Pierre et Marie Curie, 4 Place Jussieu, 75230 Paris Cedex 05, France

Pb and Sr isotopic compositions from the Indian Ocean (active ridges, old ocean floor and aseismic ridge samples) confirm the characteristic nature of the mantle record in this region. The results emphasize the importance of mixing processes between the lower mantle (oceanic-island basalt source), and the upper mantle (ridge-basalt source). The isotopic characteristics of the Indian Ocean islands seem to be in agreement with the hypothesis of the reinjection of sediments into the mantle.

SINCE the first studies of Schilling and co-workers, it has become clear that chemical and isotopic variations occur along the North Atlantic Ridge¹⁻⁴. Schilling interpreted these variations as the result of injections of lower mantle material (hotspot) into the upper mantle which is the 'usual' source of oceanic lithosphere. Langmuir *et al.*⁵ have shown that such a bipolar mixing process is insufficient to explain the observed variations and that a multipolar mixing process is necessary. A simple way⁶ of multipolar mixing is to mix two reservoirs which are themselves chemically and isotopically heterogeneous.

This kind of mixing process could explain the negative correlation between the Nd and Sr or the positive correlation between the Sr and Pb isotopic compositions of ridge basalts⁷⁻¹⁰. However, the isotopic variations observed along a ridge or within a mid-oceanic ridge basalt (MORB) population can also be interpreted in terms of local mantle heterogeneity created by various phenomena such as continental crust extraction, sediment reinjection and basalt extraction, all of them operating long before the time of ridge genesis.

It is, therefore, possible to define two contrasting models. Model 1 consists of an isotopically homogeneous and well-mixed upper mantle reservoir, which is the source of normal MORB, and of an underlying isotopically heterogeneous lower mantle reservoir, which could be the source of oceanic islands. Blobs injected from the lower mantle reservoir at the ridge crest would mix with the upper mantle reservoir and, in such conditions, it is difficult to sample non-contaminated parts of the upper mantle reservoir because ridge crest basalts are 'polluted' to various degrees. In model 2, the upper mantle reservoir is isotopically heterogeneous and this heterogeneity has been created over a long period of time by various phenomena (one of which could even be the injection of sediments in the upper mantle). Thus, when we analyse ridge basalts we observe this heterogeneity directly.

To test these models, we have analysed a series of samples from the Indian Ocean. This study was undertaken because preliminary analyses of samples from Kerguelen¹¹ and Reunion¹² yielded peculiar compositions reflected in the isotopic Sr-Pb correlation diagram. We first had to verify whether these anomalies were small-scale or regional phenomena. After verifying that Indian Ocean island anomalies were actually regional, we decided to test our two models by analysing them. In model 1, a mixing of a normal component with abnormal blobs should yield abnormal variations in ridge basalts whereas in model 2, the Indian Ocean ridge basalts should belong to the common array, particularly in the Sr-Pb correlation diagram.

In looking for the causes of the variations along the ridges, we shall also discuss various explanations of the specific characteristics of some islands of the Indian Ocean.

Results

The Sr and Pb isotopic composition of these basaltic rocks were analysed using techniques described previously^{13,14}. The results and blanks measured during this study are reported in Table 1.

(1) Oceanic Islands: We first analysed basalts from different islands including Kerguelen, Reunion, Comores, St Paul, Amsterdam and Crozet (sample locations are indicated in Fig. 1). For Kerguelen we analysed only two samples for Pb isotopes, and used data taken from the study by Dosso *et al.*¹¹. We added to this data set, the analyses of samples from the Ninetyeast Ridge. Basalts from this aseismic ridge are chemically very similar to those of the oceanic islands and we shall consider these data as a whole.

The Pb isotopic results are reported in the binary diagrams ($^{206}\text{Pb}/^{204}\text{Pb}$ plotted against $^{207}\text{Pb}/^{204}\text{Pb}$) and ($^{206}\text{Pb}/^{204}\text{Pb}$ plotted against $^{208}\text{Pb}/^{204}\text{Pb}$) on Fig. 2. For Kerguelen, Crozet, Amsterdam and St Paul islands, values of the $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{208}\text{Pb}/^{204}\text{Pb}$ ratios are larger than those of North Atlantic and Pacific islands, but are similar to those of Tristan da Cunha and Gough islands in the South Atlantic and similar to the Walvis Ridge¹⁵ (comparisons are made using Sun's compilation¹²). In

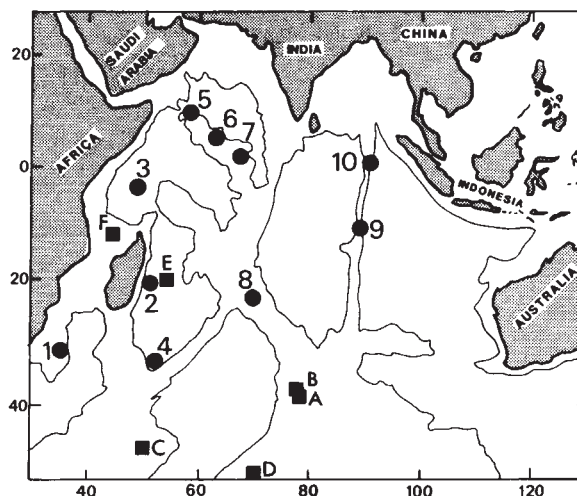


Fig. 1 Location map for the different samples from the Indian Ocean. Numbers and letters are as used in Table 1.

contrast, the Comores data do not show any $^{207}\text{Pb}/^{204}\text{Pb}$ or $^{208}\text{Pb}/^{204}\text{Pb}$ anomalies. They are very similar to data obtained on lavas from Kenya¹⁶. The isotopic data for La Reunion are intermediate between these two populations.

The Sr isotopic composition for the Indian Ocean islands appears to be very radiogenic (0.7035–0.7054) and supports the early suggestions of Hedge *et al.*¹⁷. The coupled Pb-Sr results, plotted in Fig. 2, reveal that in contrast with the North Atlantic and Pacific trend, the Indian Ocean islands define a negative trend with two endmembers: Kerguelen (for which the $^{206}\text{Pb}/^{204}\text{Pb}$ values are the lowest and the Sr values the highest) and Comores islands. This negative correlation had been suggested first by Vidal and Dosso¹⁸ who considered it to be general at the scale of the world and, from it, suggested a global extraction model with separation of a sulphide phase going into the core. This model does not agree with the North Atlantic¹² and East Pacific¹⁹ islands data which display a positive Pb-Sr correlation.

Note that in every diagram the Ninetyeast Ridge basalts behave exactly like Indian Ocean island basalts; considering the fact that they were very close to Kerguelen in the past, this result is not surprising. The isotopic identity between the Ninetyeast Ridge and the Kerguelen islands has already been noted^{20,21} for Sr.

(2) Oceanic crust basalts: We have also analysed samples collected on ancient oceanic crust during the *Glomar Challenger* (Leg 25)²² programme as well as very fresh glasses from young samples of the Indian Ridge; their location is given in Fig. 1. We also use data obtained by Cohen *et al.*¹⁰ and Sun¹². Pb isotopic values of samples from the Rodriguez triple junction and from the Carlsberg Ridge yield very low $^{206}\text{Pb}/^{204}\text{Pb}$ (17.35–18.30).

NOAA 3–10 and MO 23 DRO2 give the least radiogenic values for the $^{206}\text{Pb}/^{204}\text{Pb}$ ratio that have been observed on a ridge and plot within the field relative to the present-day geochron. If such samples were to be considered to be the most representative of the upper mantle (because they are the least contaminated by oceanic-island basalts), this reservoir should be considered as being slightly in the B field. Therefore, the extraction of continental crust should leave a residual (depleted) upper mantle with B type characteristics. This position should not be considered as abnormal as the continental crust on average is in the J field. However, these samples yield $^{208}\text{Pb}/^{204}\text{Pb}$ ratios that are larger than those obtained on other ridges for similar $^{206}\text{Pb}/^{204}\text{Pb}$ ratios. The results obtained by Cohen *et al.*¹⁰ on the Indian southeastern ridge do not show such an unusual character and plot within the normal Atlantic-Pacific trend. Therefore these characteristics seem to be specific

Table 1 Sr and Pb isotopic compositions determined for the samples studied: blanks were less than 200 pg for Sr and 0.8 ng for Pb

Sample no.		$^{206}\text{Pb}/^{204}\text{Pb}$	$^{207}\text{Pb}/^{204}\text{Pb}$	$^{208}\text{Pb}/^{204}\text{Pb}$	$^{87}\text{Sr}/^{86}\text{Sr}$	
1	249-33.3	18.133	15.537	38.53		Tholeiites
	249-33.2	18.148	15.552	38.54	0.70 361 ± 8	
2	239-21.1	17.717	15.426	37.24		
	113-119					
	Leaching	17.837	15.454	37.39	0.70 362 ± 9	
3	240-71	18.679	15.518	38.14	0.70 397 ± 9	
	126-134					
4	245-19-1	18.03	15.57	37.93		
	145-15.0	17.921	15.556	37.79	0.70 287 ± 7	
	glass					
5	NO AA 3-10	17.353	15.425	37.19	0.70 262 ± 3	Ninetyeast Ridge
	9°49.5'N 57°56.7'E					
6	NOAA 6-9	17.994	15.490	37.80		
	glass					
	3°47'N 63°52'E					
7	NOAA 9-96	17.985	15.446	37.77	0.70 279 ± 4	
	glass					
	1°39'5"N 67°46.4'E					
8	MO 23 DR 02	17.541	15.423	37.42	0.70 332 ± 5	
	glass					
	MO 23 DR 04					Islands
	glass	17.790	15.434	37.55		
9	214 481 7683	18.367	15.607	38.84	0.70 461 ± 5	
	DSDP					
10	216 22 373	17.93	15.553	38.71	0.70 543 ± 6	
	DSDP					
A	St Paul 12	18.684	15.653	39.05	0.70 358 ± 6	
B	Amsterdam	19.170	15.708	39.81	0.70 386 ± 6	
	NA 23					
C	Crozet	18.913	15.679	39.29	0.70 411 ± 8	
D	Kerg 15 G 16	18.477	15.604	39.20		
	Kerg 15 233	18.037	15.540	38.92		
E	Re 50	18.938	15.604	39.03	0.70 409 ± 3	Islands
	Re 56 Reunion	18.874	15.575	38.92	0.70 417 ± 4	
F	MI 40	19.359	15.583	39.41	0.70 369 ± 3	
	MI 159 Comores	19.249	15.515	39.19	0.70 367 ± 4	

The standard deviation for lead isotopic ratio is given as:

$$\sigma = (\sigma_{\text{mes}}^2 + R_m^2 \times \delta_m^2 \times \sigma_e^2)^{1/2}$$

σ_{mes} = statistical error of the run; σ_e = error on the mass discrimination factor estimated from the SRM 981 reproducibility; δ_m = mass difference of the measured ratio R_m . The major factor of uncertainty comes from σ_e . Typical uncertainties (1σ) are 0.012 for $^{206}\text{Pb}/^{204}\text{Pb}$, 0.014 and 0.03 for $^{207}\text{Pb}/^{204}\text{Pb}$.

to the Carlsberg Ridge. The results obtained by Cohen and O'Nions²³ on two samples located at 61° 50'E and 59°52.4'E on the Indian Ridge are very similar to ours, although small differences between the different oceans may be noted when the whole data^{12,20,23} are considered.

Results of the analysis of ancient oceanic crust samples obtained by the IPOD drilling programme are identical to those for young ridges (Site 249 and Site 245) and for typical 'normal ridge' (Site 240). Values for Site 240 (north of the Comores) are similar to those of the Comores islands.

As previously outlined²⁴, the Sr isotopic ratios are higher for the Indian Ocean ridge basalts than for those of the other oceans. This is true for raw data obtained after strong leaching experiments as well as for young or old oceanic crust, and when these results are reported on the Pb-Sr isotopic correlation diagram, they all plot slightly above the trend found for the Atlantic and Pacific oceans^{6,9,10,13}; this observation remains true even when considering the young fresh glasses only. Such a position in the Pb-Sr diagram clearly indicates that the Indian Ocean ridge samples have the lowest values for Pb isotopes, but this is not true for the $^{87}\text{Sr}/^{86}\text{Sr}$ ratios.

Discussion

The isotopic results obtained for the Indian oceanic crust are clearly different from those of the Pacific or the North Atlantic oceans and a choice between the two models proposed can be made: model 1 seems to give the best explanation of our data.

The regional isotopic variations observed for the MORB closely follow the regional variations for the Indian Ocean islands. The island source material which disturbs the MORB reservoir can also be recognized through the variations measured in the oceanic crust basalts. With the additional constraint that the oceanic island source (lower mantle?) material composition must be heterogeneous, Schilling's interpretation remains basically correct, with the result that, in global evolution models of the crust-mantle system, if oceanic-island basalts came from the lower mantle, transfer of lower mantle material into the upper mantle must be considered as a fact.

If this interpretation was correct, the ridge basalts reservoir should be isotopically rather homogeneous on a world-wide scale. However, it is perturbed by the injection of blobs which

are, in contrast, isotopically extremely heterogeneous. In this case, we can characterize the uncontaminated endmember by the intersection of such contamination lines with the Sr-Pb correlation found for the Atlantic and Pacific oceans. We thus obtain depleted mantle values: $^{87}\text{Sr}/^{86}\text{Sr} \sim 0.70200$; $^{206}\text{Pb}/^{204}\text{Pb} \sim 17.2$; $^{207}\text{Pb}/^{204}\text{Pb} \sim 15.4$; $^{208}\text{Pb}/^{204}\text{Pb} \sim 36.8$. Using the Nd-Sr correlation^{7,8,25}, these would correspond to an ϵ_{Nd} of +13. These values should be considered as the best estimate of the uncontaminated (upper mantle) composition.

The remaining, and much more difficult, problem is the cause of the isotopic heterogeneity of oceanic islands and, especially, that of the Indian Ocean. First, we should recall that in the Sr-Pb diagram, the Indian Ocean island data define a negative trend, and second, that in the Pb-Pb isotopic diagram, these islands define a trend roughly parallel to the usual MORB correlation but shifted towards higher $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{208}\text{Pb}/^{204}\text{Pb}$ values and, finally, that in the Nd-Sr diagram, the trend coincides with the mantle array although some values are higher than admitted for the bulk Earth²⁶. We have no definitive answer to this problem and the following should be considered as a tentative explanation. We will thus consider the existence of two reservoirs which may be identified with the upper and the lower mantle.

Model A: lower mantle with pristine and depleted domains. The lower mantle is the source of blobs, but it is not convecting as a whole and therefore it is extremely heterogeneous. Parts of that lower mantle, that have an almost pristine composition for all elements, are the source of the Kerguelen volcanism. The $^{87}\text{Sr}/^{86}\text{Sr}$ (0.705) value is almost planetary whereas those for lead isotopes correspond to closed system values. Other parts of the lower mantle are extremely enriched in U/Pb as, for example, beneath the Atlantic. The Indian Ocean islands originate from sources which are a mixture of these two types of lower mantle.

Model B: contamination by continental crust material. Indian Ocean islands are abnormal because they have been contaminated by continental crust material. Numerous studies have shown (see the compilation of Doe and Zartman²⁷) that the continental crust isotopic characteristics are different from those of the oceanic islands. Despite their great variety of isotopic compositions, nearly all the continental crust samples have radiogenic values for $^{87}\text{Sr}/^{86}\text{Sr}$, $^{208}\text{Pb}/^{204}\text{Pb}$ and $^{207}\text{Pb}/^{204}\text{Pb}$. Contamination may have occurred locally because one may consider that the Kerguelen or the Ninetyeast Ridge are remnants of ancient continental crust. However, the geophysical data do not give a definitive answer to the problem of the nature of the basement underlying Kerguelen or the Ninetyeast Ridge²⁸. Qualitatively the ^{87}Sr , ^{208}Pb and ^{207}Pb enrichment of these oceanic islands would suggest some kind of contamination by the continental crust.

Another way of contaminating the oceanic islands source would be the reinjection of sediments into the mantle. This has been already proposed^{27,29,30} to explain the oceanic islands isotopic characteristics. This reinjection would create an enriched domain in the lower part of the upper mantle (or in the lower mantle) and this enriched domain may contaminate the hotspots when they rise towards the surface. (Here, we consider the fact that basalts from islands are more contaminated than basalts from ridges.)

Figure 3 shows data for oceanic islands throughout the world and reveals that it is possible to define large-scale areas for which Nd-Sr-Pb isotopic compositions are roughly the same. The South Atlantic and the southern Indian oceans are quite different from the North Atlantic and the eastern Pacific oceans. This geographical mapping suggests that it is possible to define large-scale domains for the oceanic island source. More data are needed before the complex isotope chemistry of the Earth's mantle may be understood.

This isotopic similarity between the values of Tristan da Cunha, Gough and some of the islands of the Indian Ocean allows us to use the ^3He results of Kurz *et al.*³¹ for these islands; their results favour model B.

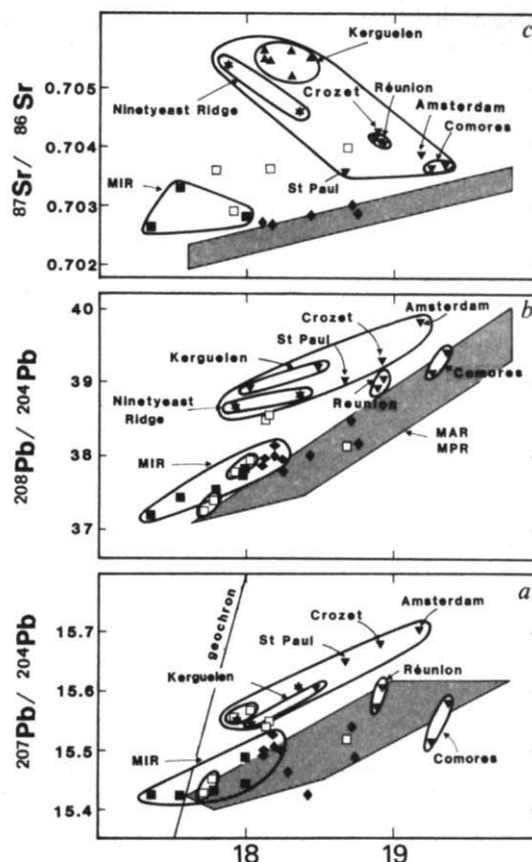


Fig. 2 Results for samples from Indian Ocean. *a*, $^{207}\text{Pb}/^{204}\text{Pb}$ plotted against $^{206}\text{Pb}/^{204}\text{Pb}$. ■, Samples from the Indian Ridge; □, old oceanic crust; ♦, results obtained by Cohen *et al.*^{10,23}; ▼, results obtained by Sun¹²; ★, Ninetyeast Ridge. The left side of the geochron is labelled as B field, and the right side as J field. *b*, $^{208}\text{Pb}/^{204}\text{Pb}$ plotted against $^{206}\text{Pb}/^{204}\text{Pb}$. *c*, $^{87}\text{Sr}/^{86}\text{Sr}$ plotted against $^{206}\text{Pb}/^{204}\text{Pb}$. ▲, Data obtained by Dosso *et al.*¹¹. The hatched area is the domain of MORB values for the Atlantic and Pacific oceans (refs 4, 9, 10, 12, 26 and B. Hamelin, B.D. and C.J.A., in preparation).

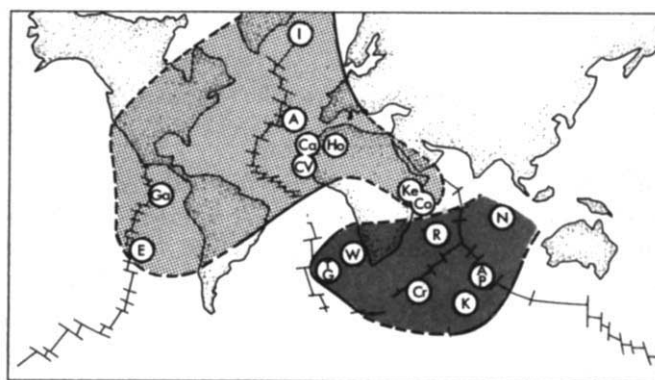


Fig. 3 General map indicating the large geographical domains with similar isotopic characteristics. I, Iceland; A, Azores; H, Hoggar; Ca, Canaries; Cv, Cape Verde; F, Fernando de Noronha; Ga, Galapagos Islands; E, Easter Island; Co, Comores; T, Tristan da Cunha; G, Gough; W, Walvis Ridge; Cr, Crozet; R, Reunion; K, Kerguelen; A, Amsterdam; P, St Paul; N, Ninetyeast Ridge; Ke, Kenya. Data are from refs 11, 12, 16, 19, 32 and this work.

Indeed, a very important difference between He isotopic values and those of the other radiogenic tracers Sr, Pb and Nd can be found in their position relative to the main endmembers considered in present models (the MORB source mantle, the virgin primitive mantle, the continental crust and sediments). For Sr, Nd and Pb, the virgin mantle has values intermediate between those for the residual mantle (MORB source) and those for the continental crust and sediments. Thus, values

similar to those of the virgin mantle may be interpreted either as actually originating from 'primitive' portions of the mantle, or resulting from the mixing of the two extreme reservoirs (MORB and sediments). For He in contrast, the virgin mantle values are not intermediate between those for the two other endmembers. This results from the fact that the complementary reservoir of the mantle for the rare gases is not the continental crust but the atmosphere (in the case of He the atmosphere is a leaky reservoir). This difference will thus enable a choice to be made between the various models.

The $^3\text{He}/^4\text{He}$ values for Gough and Tristan da Cunha are more radiogenic than the MORB values³¹, themselves being more radiogenic than the pristine value. These results thus exclude the primitive mantle interpretation for these islands, as well as for the Kerguelen islands, which have similar Sr, Nd and Pb characteristics. In contrast, mixing with sediments may explain all the isotopic data.

Conclusions

Some Pb–Sr isotopic results obtained on islands of the Indian Ocean are similar to those of the North Atlantic Ocean (Comores, for example) and others are similar to those of the South Atlantic (Kerguelen, for example). These Pb–Sr–Nd isotopic similarities on a large geographical scale suggest a broad cartography of large domains with specific characteristics: North Atlantic and West Pacific on the one hand, and South Atlantic and Indian Ocean on the other.

The Pb, Nd, Sr and He characteristics of the endmembers represented by Gough, Tristan da Cunha and Kerguelen are interpreted by the mixing of lower mantle material with subducted sediments.

The isotopic peculiarity of the Indian Ocean islands is also found, with less intensity, in the Indian Ridge tholeiites. This confirms the hypothesis of mixing between the oceanic islands source and the ridges tholeiites source.

We thank J. L. Cheminee, R. Schlich, R. Hekinian and M. Kurz for providing samples, B. Hamelin and C. Gariepy for constructive criticism, and D. Rousseau for experimental assistance. IGP contribution no. 628.

Received 7 October 1982; accepted 23 February 1983.

- Schilling, J. *Nature* **242**, 565 (1973).
- Hart, S. R., Schilling, J. G. & Powell, J. R. *Nature, phys. Sci.* **246**, 104–107 (1973).
- White, M., Schilling, J. G. & Hart, S. R. *Nature* **263**, 659–663 (1976).
- Sun, S. S., Tatsumoto, M., & Schilling, J. G. *Science* **190**, 143–147 (1975).
- Langmuir, C. H., Vocke, J. R. D., Hanson, G. N. & Hart, S. K. *Earth planet. Sci. Lett.* **18**, 380–392 (1978).
- Allègre, C. J., Brevart, O., Dupre, B. & Minster, J. F. *Phil. Trans. R. Soc. A* **297**, 447–477 (1980).
- Richard, P., Shimizu, N. & Allègre, C. J. *Earth planet. Sci. Lett.* **31**, 269–278 (1976).
- De Paolo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* **3**, 249–252 (1976).
- Dupre, B. & Allègre, C. J. *Nature* **286**, 17–22 (1980).
- Cohen, R. S., Evensen, N. M., Hamilton, P. J. & O'Nions, R. K. *Nature* **283**, 149 (1980).
- Dosso, L. *et al. Earth planet. Sci. Lett.* **43**, 46–60 (1979).
- Sun, S. S. *Phil. Trans. R. Soc. A* **297**, 409–445 (1980).
- Birck, J. L., & Allègre, C. J. *Earth planet. Sci. Lett.* **39**, 37–51 (1978).
- Manhes, G., Minster, J. F. & Allègre, C. J. *Earth planet. Sci. Lett.* **39**, 269–278 (1978).
- Richardson, S. H., Erlank, A. J., Duncan, A. R. & Reid, D. L. *Earth planet. Sci. Lett.* (submitted).
- Norri, M. J. *et al. Phil. Trans. R. Soc. A* **297**, 259–271 (1980).
- Hedge, C. E., Watkins, N. D., Hildreth, R. A. & Doering, N. P. *Earth planet. Sci. Lett.* **29–34** (1973).
- Vidal, P. & Dosso, L. *Geophys. Res. Lett.* **5**, 169–172 (1978).
- White, M. *Yb Carnegie Instn Wash.* **78**, 331–335 (1979).
- Reddy, V. V., Subbarao, K. V., Reddy, G. R., Matsuda, G. R. & Hekinian, R. *Mar. Geol.* **26**, 99–117 (1978).
- Whitford, D. J. & Duncan, R. A. *Yb Carnegie Instn Wash.* **77**, 606–613 (1978).
- Schlich, R., Simpson, E. S. & Vallier, T. L. *Int. Rep. D.S.D.P. Leg 21*, 743–579.
- Cohen, R. S. & O'Nions, R. K. *J. Petrol.* **23.3**, 299–324 (1982).
- Subbarao, K. V. & Hedge, C. E. *Earth planet. Sci. Lett.* **18**, 223–228 (1973).
- O'Nions, R. K., Hamilton, P. J. & Evensen, N. M. *Earth planet. Sci. Lett.* **39**, 13–22 (1977).
- Allègre, C. J., Othman, B., Polve, M. & Richard, P. *Phys. Earth planet. Sci. Inter.* **19**, 293–306 (1979).
- Doe, B. R. & Zartman, R. E. in *Geochemistry of Hydrothermal Arc Deposits* (ed. Barnes, H. L.) 22–66 (Wiley, New York, 1979).
- Ben Avraham, Z., Nur, A., Jones, D. & Cox, A. *Science* **213**, 47–54 (1981).
- Armstrong, R. L. *Rev. Geophys.* **6**, 175–199 (1968).
- O'Nions, R. K. & Hamilton, P. J. *J. geophys. Res.* **84**, 101–6091 (1979).
- Kurz, M. D., Jenkins, W. J. & Hart, S. R. *Nature* **297**, 43–46 (1982).
- Allègre, C. J., Dupre, B., Lambret, B. & Richard, P. *Earth planet. Sci. Lett.* **52**, 85–92 (1981).

Small-angle X-ray scattering from myosin heads in relaxed and rigor frog skeletal muscles

F. R. Poulsen & J. Lowy

The Open University, Oxford Research Unit, Foxcombe Hall, Berkeley Road, Boars Hill, Oxford OX1 5HR, UK

Low-angle X-ray diffraction patterns from relaxed and non-overlap rigor muscles show a central region of diffuse scattering (disk) which is circularly symmetrical, behaves as solution scattering and comes predominantly from myosin heads. In full-overlap rigor the disk is compressed in the diagonal direction, indicating that the myosin heads have a bent shape and a preferred orientation consistent with a 45° angle of attachment to actin.

DURING muscular contraction, interdigitating arrays of actin and myosin filaments slide further into each other^{1,2}. The required energy release takes place between the two types of filament and involves a cyclical action of 'cross-bridges'³. The central problem is to understand the relationship between chemical, structural and mechanical aspects of the cross-bridge cycle. Here we describe a new approach to the study of the structural behaviour of the cross-bridges.

Cross-bridges can be seen in the electron microscope as projections from the surface of the myosin filaments^{4–6}. They are largely composed of the S-1 and S-2 subfragments of the myosin molecule, two globular S-1 heads being joined to one double α -helical S-2 rod⁷. The latter connects the heads to the double α -helical light meromyosin rods which in vertebrate striated muscle form the backbone of the filament and constitute the third major component of the myosin molecule⁷.

X-ray diffraction has been extensively used to obtain structural information about the cross-bridge cycle mainly because it can be applied to the living system. Experience from crystallography naturally led to the choice of muscles such as the striated frog sartorius which produce relatively sharp peaks⁸. However, a full interpretation presented difficulties which, besides the usual cylindrical mixing of intensity about the fibre axis, arose partly from lattice sampling on the meridional peaks⁹ and partly from the presence of various types of disorder¹⁰. Consequently, even the configuration of the projections in the relaxed muscle is not fully understood. Nevertheless, recent time-resolved studies of the sharp peaks during contraction have given very interesting results^{11–13}. Thus, it was shown that myosin heads move towards actin slightly ahead of tension rise, whilst their three-dimensional order breaks down and their axial repeat becomes more prominent; furthermore, rapid

length changes imposed during contraction were found to cause the axial repeat of myosin heads to become less prominent and this was interpreted to indicate that they were attached to actin.

In spite of this progress, it has not so far been possible to determine numbers of attached heads and their spatial behaviour, the kind of structural information required for an understanding of the cross-bridge cycle. The main reason is that the sharp peaks which contain the spatial information for the resting state practically disappear during contraction and no new peaks can be detected^{8,13}. Accordingly, some of the lost information should show up as diffuse scattering. It is well known that even in the resting muscle a very strong level of diffuse scattering is present superimposed on the sharp peaks. If a significant part of this scattering were due to myosin heads it would contain structural information about them for both resting and active muscle. In previous work with frog muscles the diffuse scattering was not associated with the contractile apparatus. It was therefore treated as background noise and often artificially reduced in the process of printing the pattern from the X-ray film^{8,14,15}.

We conclude from the X-ray results presented here that the diffuse scattering is, in fact, predominantly due to myosin heads. This reveals a situation so far unique for molecules in a biological system. The light meromyosin filament framework, rather than imposing a strict crystalline order on the heads protruding from it, maintains them in conditions where they produce scattering of a kind which, dependent on the accuracy of the experimental data, can provide information about numbers, size, shape and orientation of myosin heads during the cross-bridge cycle in living muscle.

X-ray scattering patterns

Using frog sartorius and semitendinosus muscles we find that in the relaxed state at both short and long sarcomere lengths, X-ray patterns show a strong central region of diffuse scattering (disk) superimposed on the layer line system (Fig. 1*a, c*). When relaxed muscles are stretched the layer lines become weaker¹⁵ and the disk stronger. Densitometry shows that the disk has circular symmetry (Fig. 2*a, c*).

If rigor is induced in muscles at sarcomere lengths where there is full overlap between the actin and myosin filaments, the relaxed-state myosin layer lines disappear and are replaced by rigor layer lines due to myosin heads attached to actin⁸. At the same time the intensity of the disk is greatly reduced (Figs 1*b, 2b*) and its shape becomes asymmetrical (Fig. 3*b*).

In semitendinosus muscles in which rigor is induced at non-overlap length, the relaxed-state myosin layer lines disappear but they are not replaced by any new layer lines^{8,15}. At the same time the intensity of the disk is greatly enhanced (Figs 1, 2*d*), and now the disk has retained its circular symmetry (Fig. 3*a*).

Structural interpretations

We interpret our results to show that the disk is predominantly due to myosin heads which scatter in relaxed and non-overlap rigor states as if they were in solution and in full-overlap rigor as if they had a preferred orientation.

Two obvious conditions for the solution behaviour are that the scattering must be circularly symmetrical and that its profile must be very similar to that obtained from X-ray scattering experiments with S-1 subunits in solution^{16,17}. Figures 2*a, c, d* and 3*a* show that these conditions are indeed fulfilled. The deviation close to the equator ($l < 0.05 \text{ nm}^{-1}$) is explained in Fig. 3 legend. The similarity between the two data sets is also revealed by calculating (using procedures described in Fig. 4*a* legend) the radius of gyration, the molecular volume and the length of the mean chord. The results are, respectively, 3.1 nm, 150 nm³ and 4.0 nm and these compare well with the corresponding values determined from solution scattering studies of S-1, namely 3.28 nm, 151 nm³ and 4.25 nm (refs 16, 17).

We next consider the difference in scattering between non-overlap and full-overlap rigor. As the chemical conditions are

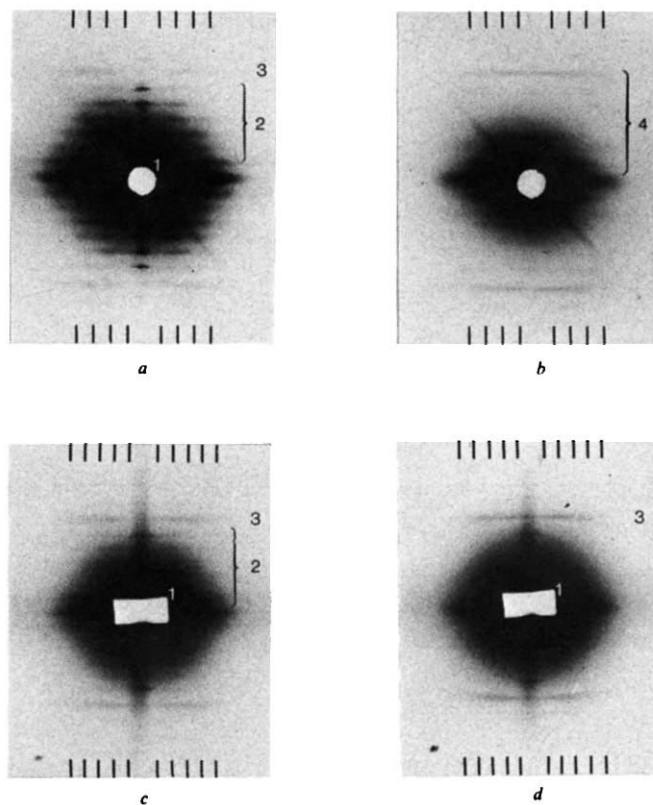
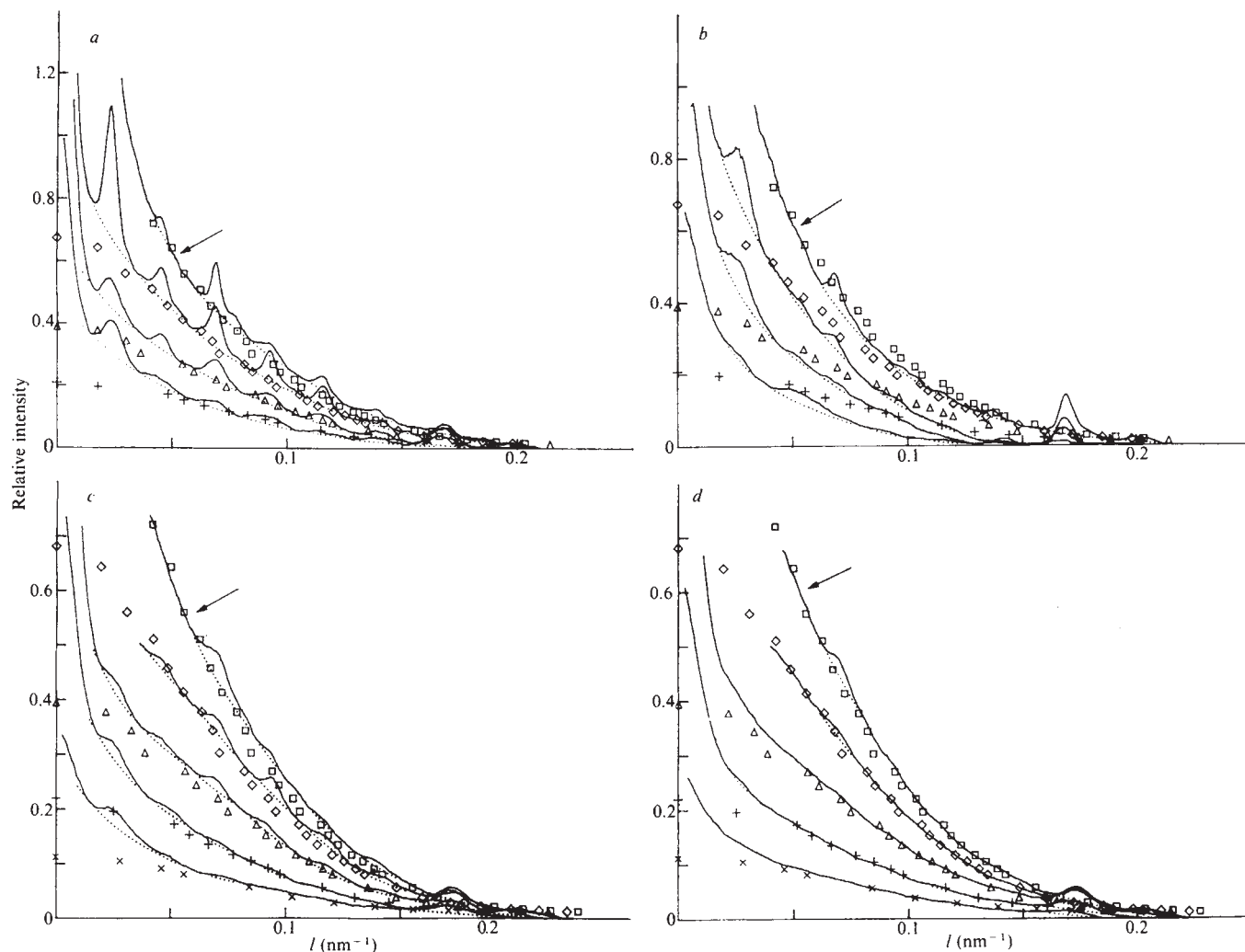


Fig. 1 Low angle X-ray patterns from muscles shown so that they correspond to the fibre axis being vertical. The X-ray methods used were similar to those described previously⁸. Frog muscles were mounted isometrically and maintained in Ringer's solution³¹ at about 2 °C in Perspex cells with mylar windows for the passage of X-rays. Rigor muscles were produced by iodoacetate poisoning⁸. The exposure times were about 14 h. Within each experiment (same muscle and camera) the exact exposure time was used to scale the intensity data together. The vertical lines indicate where these and similar frog patterns were scanned to produce densitometer tracings as shown in Fig. 2. *a*, Living relaxed frog sartorius at sarcomere length 2.3 μm (full overlap between actin and myosin filaments). Note the disk of diffuse scattering (1) and the layer lines due to myosin (2) and actin (3). Because of the presence of the myosin layer lines it is difficult to assess the exact shape of the diffuse disk by eye, but densitometry (Fig. 2*a*) shows that it is practically circularly symmetrical. *b*, Same muscle as in *a* after induction of rigor at sarcomere length 2.3 μm . The myosin layer lines have been replaced by the rigor layer lines (4) due to myosin heads attached to actin monomers⁸. The intensity of the diffuse scattering has greatly decreased and its shape has become compressed in the diagonal direction (Fig. 3*b*). *c*, Living relaxed frog semitendinosus stretched to sarcomere length 4.2 μm (non-overlap). Note the disk of diffuse scattering (1) and the layer lines due to myosin (2) and actin (3). *d*, Same muscle as in *c*, rigor induced at sarcomere length 4.2 μm . The layer lines due to the regular arrangement of myosin projections in the relaxed state have disappeared but no rigor layer lines have become visible¹⁵ (compare with *b*). The diffuse scattering (1) has increased by about 50% (Fig. 2*d*) and has retained its circular symmetry, which in this pattern is obvious when the equatorial diffraction is blocked out. This is confirmed by densitometry (Figs 2*d*, 3*a*). For the purpose of recording X-ray scattering our cameras have practically a point focus. The S-1 solution scattering data¹⁶ with which we compare our data (Figs 2, 3) were obtained with a line focusing camera. We have confirmed that this makes no difference for the compared range of data by Fourier transforming the distance distribution curve¹⁶ which is corrected for slit width.

the same the difference must lie solely in the arrangement of the myosin heads which is caused by their attachment to actin in full-overlap rigor. This immediately shows that at least 85% of the scattering comes from myosin heads (Fig. 3*a, b*). Most probably the same applies to the remaining 15%. This follows



from the deduction that because of its asymmetrical component (absent in non-overlap rigor), the scattering remaining in full-overlap rigor must also be dominated by myosin heads; Fig. 3 legend gives further details.

Lastly, there is evidence from the similarity between the diffuse scattering in the relaxed state and non-overlap rigor. As a scale factor in intensity is the only difference it seems extremely unlikely that the diffuse scattering in the relaxed state could be due to structures other than myosin heads. There are further implications from the full-overlap rigor results. Thus, the decreased intensity of the scattering indicates that the myosin heads have gained more order, and the asymmetry of the scattering that they have a preferred orientation.

The diagonal compression of the diffuse intensity distribution (Fig. 3b) seems at first sight puzzling. If the attached myosin head is modelled as a straight rod attached at an angle of 45° , exactly the opposite result is expected (this angle has been suggested from electron microscope evidence from insect flight muscle in rigor¹⁸). Our data may be explained on a rod-shaped head model if one assumes that some heads are oriented in parallel and some at right angles to the fibre axis. However, this explanation disagrees with saturation transfer electron paramagnetic resonance¹⁹ and linear dichroism²⁰ data which show that in the rigor state myosin heads are attached to actin with a very narrow range of angles. The answer must therefore be sought in the shape of the myosin head which, rather than being rod-like, is likely to be significantly bent, as has been deduced from both electron microscopy and X-ray scattering experiments¹⁶. Our present results are insufficient to determine the shape of the head, but they impose some interesting restrictions.

The most effective way of explaining the observed asymmetry in the intensity distribution is on a model in which the mass of the myosin head is distributed into two parts, one plate-shaped and the other rod-shaped. The plate part (whose normal should be parallel to the fibre axis) gives a large contribution only on the meridian. The rod part (which should also be parallel to the fibre axis) makes a large contribution only on the equator. These two parts can be combined in a variety of ways which correspond to different orientations of the myosin head (direction of the longest chord), including one at an angle of 45° to the fibre axis. With their assigned orientations, no matter how the plate and the rod part are combined, the result will always be a highly bent structure. We have confirmed by extensive model calculations that the diagonal compression can be explained by bent shapes of the myosin head with a 45° angle of attachment. Our data do not prove that the attachment angle is 45° but they are consistent with that orientation.

Summarizing and also taking account of previous work, our structural interpretation of the diffuse scattering is as follows. In the relaxed state, myosin heads have both an ordered aspect giving rise to layer lines and, because they are free to move, a disordered aspect which is responsible for the scattering. If rigor is induced in the muscle at full-overlap the myosin heads attach to actin^{8,15} with a preferred orientation¹⁹⁻²¹, causing the diffuse scatter to decrease and adopt an asymmetrical shape. At the same time the relaxed-state myosin layer lines are replaced by rigor layer lines^{8,15}. If rigor is induced at non-overlap, the myosin heads lose their ordered relaxed-state aspect, as evidenced by the disappearance of the layer lines, but being unable to attach to actin they do not assume any new order and so the scattering increases.

Fig. 2 Densitometer tracings of X-ray patterns similar to those shown in Fig. 1, and comparisons of these data with S-1 solution scattering data¹⁶. Because of the cylindrical mixing of the intensity about the fibre axis we describe our data in terms of cylindrical coordinates. The film records a two-dimensional section through reciprocal space and in general four symmetry-related points are characterized by their axial (l) and radial (R) coordinates. The distance from the origin $s = \sqrt{l^2 + R^2}$ is related to the scattering angle (2θ) and the X-ray wavelength (λ) by $s = 2 \sin(\theta)/\lambda$. We have used this relationship in transposing the solution scattering data published¹⁶ in the form of $\log(I(4\pi R_g \sin(\theta)/\lambda))$, where $R_g = 3.28$ nm, into $I(l, R)$ to compare them with our muscle data which, as indicated in Fig. 1, were scanned in the axial direction at various radial positions. The data were recorded using a modified Joyce-Loebl microdensitometer MK III C which, linked to a computer, scans the film pattern in two dimensions. The main steps in the data treatment were as follows. The measured intensities were calibrated with respect to exposure time and averaged over the four symmetry-related quadrants. The background (which was assumed to be constant for each axial scan) was taken to be equal to the averaged intensity at an s value of about 0.22 nm^{-1} and then subtracted from the intensity distributions. This relatively simple background correction was considered adequate for our present purpose, first, because at the chosen s value the intensities of both the S-1 solution scattering data and the muscle data are less than 1% of their maximal values, and second, because recordings of the scattering from the camera and the bathing solution alone show that in the region where the data are used, this background varies very slowly in the axial direction and is negligibly small compared with the scattering produced when the muscle is present. *a*, Pattern from living relaxed semitendinosus muscle at sarcomere length $2.7 \mu\text{m}$ recorded on a 360-mm mirror-crystal camera. *b*, Pattern from the same muscle as in *a*, with rigor induced at the same sarcomere length of $2.7 \mu\text{m}$ recorded on the same camera. *c*, Pattern from living relaxed semitendinosus muscle at sarcomere length $4.2 \mu\text{m}$ recorded on a 170-mm mirror-crystal camera. *d*, Pattern from same muscle as in *c*, with rigor induced at the same sarcomere length of $4.2 \mu\text{m}$ recorded on the same camera. As regards the extent of overlap in rigor muscles, it is known that when myosin heads attach to actin, the 5.9-nm layer line shifts its maximum towards the meridian⁸. We have used this shift as a control for the amount of overlap deduced from laser light diffraction because that value may be in error due to a possible spread in sarcomere lengths. Densitometry shows that in *d* there is no shift in the maximum of the 5.9-nm layer line, thus demonstrating that this pattern was obtained from a part of the muscle at non-overlap. Supporting this is the absence of rigor layer lines which can be established by visual inspection (Fig. 1*d*). Densitometry of the patterns in *a* and *b* shows that here the maximum of the 5.9-nm layer line shifts from $R = 0.08$ to 0.05 nm^{-1} which we find is characteristic of what happens when a muscle goes into rigor at full overlap. Furthermore, rigor layer lines are clearly visible in Fig. 1*b*. In *a*, axial intensity distributions (—) and corresponding S-1 solution data¹⁶ are shown at R equal to 0.0270 (□), 0.0541 (◇), 0.0811 (△) and 0.108 nm^{-1} (+). The diffuse muscle scatter under the layer line peaks was defined by drawing smooth curves (· · ·) through the minima between the peaks. The muscle data were scaled to the solution data at the location indicated by the arrow. Note that, except close to the equator ($l < 0.05 \text{ nm}^{-1}$), where an agreement is not expected (for reasons explained in Fig. 3 legend), the muscle data agree very well with the solution data for all values of R . This also means that the scattering has circular symmetry. In *b*, axial distributions of total intensity (—), interpolated diffuse scattering (· · ·) and the corresponding S-1 solution data¹⁶ are shown at the same R values as in *a*. We find that relative to the relaxed state, the intensity of the diffuse scattering at the point of scaling (arrow) has decreased by a factor of 1.8. Furthermore, the shape of the intensity distribution has changed so that it no longer agrees with the solution scattering data. At an axial position (l) of about 0.05 nm^{-1} there is a fair agreement at all values of R , but with increasing l the intensity falls clearly below the solution data, especially at large values of R . This means that the shape of the diffuse scattering is compressed in the diagonal direction (see Fig. 3*b*). In *c*, axial intensity distributions (—), interpolated diffuse scattering (· · ·) and corresponding S-1 solution data are shown at $R = 0.0267$ (□), 0.0534 (◇), 0.0801 (△), 0.107 (+) and 0.134 nm^{-1} (×). The arrow indicates where the muscle data were scaled to the solution data. The same comments apply as were made for *a*. In *d*, axial intensity distributions (— and · · ·) and corresponding S-1 solution data are shown at same R values as in *c*. We find that relative to the relaxed state, the intensity of the diffuse scattering has increased by a factor of about 1.5. However, the shape of the intensity distribution has not changed and except close to the equatorial diffraction it agrees everywhere very well with the solution data. Thus, the intensity distribution has circular symmetry (Fig. 3*a*).

Discussion

Our conclusion that disordered myosin heads cause most of the scattering implies that very little is produced by anything else in the cell. As regards the contractile apparatus, the thin filaments and backbone of the thick filaments are such well ordered systems that their scattering must be negligible. There may be some scattering from the thin rod-like S-2 subunits but because their diameter is only 2 nm (ref. 7), that scatter will have a relatively flat profile and will be subtracted in our method of background correction (Fig. 2). There remain the soluble sarcoplasmic proteins whose total mass is about 50% greater than that of the myosin heads²². Moreover, glyceraldehyde phosphate dehydrogenase, which is present in the highest quantity²², has a molecular weight somewhat greater than that of the myosin head. Nevertheless, because of their very high concentration in the extrafibrillar space and their great diversity in size²², the sarcoplasmic proteins are likely to produce much less scattering.

Scopes²² has argued that most of the sarcoplasmic proteins are so large that, even if of spherical shape, they could only just be fitted between the filaments where they would severely restrict the efficiency of the sliding process. Scopes²² therefore concludes that these proteins are located in the extra-fibrillar spaces where he calculates their concentration to be 25–30% by mass. If glycogen²³ is also taken into account the value comes to 40% by mass. At such high concentrations there will be very strong scattering interference between particles. This has the effect of greatly reducing the intensity at small angles and because of the wide distribution of particle sizes (corresponding to molecular weights of 18,000–360,000)²² the resultant

intensity distribution will be much broader than that due to scatter from myosin heads. The latter constitute a practically monodisperse system of concentration about 5% by mass²² and due to their attachment to the filament backbone are evidently in a very favourable condition to produce scattering that can give useful information about the heads. We conclude that in the region where the myosin heads scatter, the sarcoplasmic proteins contribute very little; and that the scattering which they do produce has a relatively flat profile which is subtracted in our method of background correction. Scopes' interpretation²² is supported by our scattering data. If all the sarcoplasmic proteins were evenly distributed their overall concentration would be about 25% by mass. This would produce either a very flat total scattering profile, or a very low scattering intensity at small angles. Neither is seen in our patterns. The location of the soluble proteins outside the spaces between the filaments also accounts for the observation that the refractive index of the sarcoplasm is higher than even that of the A bands in isolated frog fibres².

Having dealt with the problem of scattering from other sources, we are now in a position to suggest a model which explains how the myosin heads produce the observed diffuse scattering. First, they must have large, random displacements in all directions from their average arrangement. Second, to produce circularly symmetrical scatter such elongated heads must on average be practically randomly oriented. Based on these requirements, model calculations (Fig. 4*b*) show, as expected, that the intensity decreases with decreasing distance (s) from the centre of the pattern where it becomes zero. The solution scattering curve is approached with increasing s but there is disagreement at small angles ($s = 0.004 \text{ nm}^{-1}$) which disappears

when a r.m.s. displacement of at least 9 nm is assumed. The large 9 nm r.m.s. displacement means that the intensity of the relaxed-state myosin layer lines would diminish with increasing s much faster than is observed (Fig. 1a). To account for the simultaneous presence of the diffuse scatter and the layer lines, an obvious explanation is that there are two populations of myosin heads, one with small displacements responsible for the layer lines and one with large displacements which produces the scattering. This is also necessary to explain how the intensity of the scattering can change by a factor of 1.5 without any

shape change (Figs 2d, 4a). An increase in diffuse scattering (Figs 1d, 2d, 4a) cannot be explained by an increase in the r.m.s. displacement of one population of myosin heads but is accounted for by a transfer of myosin heads from the ordered to the disordered population. Assuming that in the non-overlap rigor state all myosin heads are disordered, we calculate that about 50% of the myosin heads have the large displacements in the non-overlap relaxed state (Fig. 4a).

We suggest three possible causes for the two populations of myosin heads in the resting muscle. One is an interaction

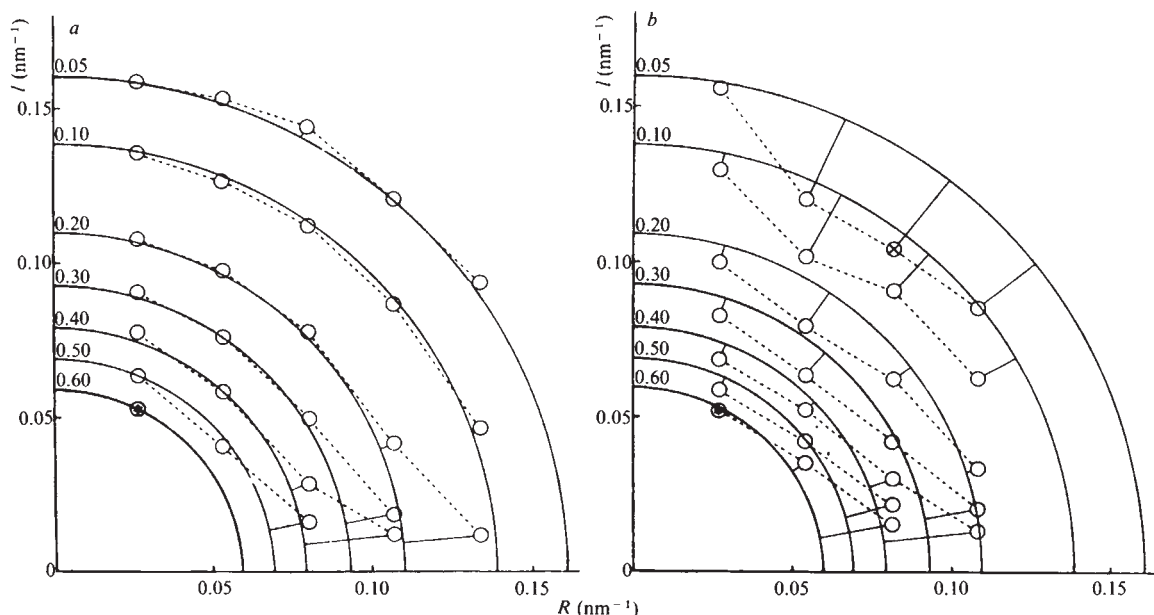


Fig. 3 Shapes of diffuse scattering are shown in terms of lines of equal intensity in l - R plots. *a, b*, Comparisons of muscle and solution scattering data. The continuous circular lines labelled with their respective intensity values represent the S-1 solution scattering data¹⁶. The muscle scattering data (○) are connected together by broken lines and to the solution data of the same intensity by solid lines pointing towards the centre. The latter lines are then a measure of the difference in shape of the two intensity distributions. The point at which the muscle data were scaled to the solution data is indicated by an asterisk. The muscle data in *a* and *b* come from Fig. 2d and *b*, respectively, abstracting from the layer line peaks along the dotted lines. Note that in *a* (non-overlap rigor) the muscle data have practically circular symmetry and in all other respects follow the solution data, except close to the equator (see below). In contrast, in *b* (full-overlap rigor) the muscle scattering is clearly compressed along the diagonal direction. In *b* the intensity of the muscle scattering at the scaling point (*) is only about 30% of that in *a*. This can be calculated from the increase by the factor 1.5 (Fig. 2d) and the decrease by the factor 1.8 (Fig. 2b) on going from the relaxed to the rigor states at non- and full-overlap, respectively; and from an increase by a factor of 1.3 on going from full- to non-overlap in the relaxed state (data not illustrated). In the diagonal direction (at ⊗ in *b*) there is less than 15% of the intensity left. As explained in the text the difference between *a* and *b* must be due to scattering from myosin heads. Because the scattering in *a* is circularly symmetrical, at least 85% of it must have come from myosin heads. Is it possible that a circularly symmetrical component of 15% in *a* could be due to structures other than myosin heads? If so, it would still be present in *b*, but this would mean that the remaining asymmetrical component of *b*—which must be due to myosin heads since it was not present in *a*—reaches zero at ⊗. This is extremely unlikely even for a population of myosin heads with a preferred orientation. Thus, we conclude that most of the diffuse scattering in *b* must come from myosin heads and so must therefore the major part of that in *a*. As regards the deviation close to the equator, this is expected on two grounds. First, the equatorial diffraction is very strong and dominates the pattern out to about $l = 0.02 \text{ nm}^{-1}$. Second, there will be an increase in intensity due to interference between the two myosin heads in the molecule. This is supported by our results from an unstriated muscle, the anterior byssus retractor of *Mytilus*. Here we have observed the same kind of effect but to a much larger extent in that the diffuse scattering is clearly separated into an inner strong and an outer weaker disk by a ring of relatively high intensity at a spacing of about 14 nm. We find that the intensity distribution of the outer disk is the same as that for S-1 in solution. Evidence will be published elsewhere that the inner disk and the ring in the anterior byssus retractor can be explained in terms of an interference between myosin heads of one or two molecules. Having dealt with the shape of the intensity distribution we now show how the general level of intensity is related to the number of attached myosin heads in the full-overlap rigor state. In the following, S denotes the scattering vector which defines a position in reciprocal space and whose magnitude is s . $f(S)$ means the Fourier transform of the structure of the myosin head and $\langle f^2(S) \rangle$ means the spherical average of its square which is equal to the intensity distribution expected from myosin heads in solution. The unattached heads would contribute with $(1-x)\langle f^2(S) \rangle$ to the intensity, where x is the fraction of heads that are attached. As explained in the text, the contribution from the latter is due to substitution disorder. Here the contribution²² amounts to: $q(1-q)f^2(S)$, where q is the fraction of the actin monomers with myosin heads attached. q is equal to xQ , where Q is the ratio between the total number of myosin heads and the total number of actin monomers in the overlap region. This ratio comes to 0.577, assuming that the frog myosin filament is three-stranded⁹. The total intensity of the diffuse scattering is then given by: $I(S) = (1-x)\langle f^2(S) \rangle + (xQ(1-xQ))f^2(S)$ and approximating $f^2(S)$ by $\langle f^2(S) \rangle$ we have: $I(s) = (1-0.423x-0.333x^2)\langle f^2(S) \rangle$. If no myosin heads were attached (that is, $x = 0$), the intensity would be $\langle f^2(S) \rangle$. If all myosin heads were attached (that is, $x = 1$), the intensity would be $0.24 \langle f^2(S) \rangle$. To obtain the observed intensity, namely, $0.30 \langle f^2(S) \rangle$, x must be 0.95. Here it is relevant to compare the situation just described in the frog muscle with that in insect flight muscle where it is known that only about 70% of the heads are attached in full-overlap rigor²⁷. This means that a large population of unattached heads will produce strong circularly symmetrical scattering. It is also known that in the insect muscle the attachment of heads follows a specific marking scheme²⁶ which has a fairly high degree of order. This means that the attached heads will diffract strongly on layer lines but will produce only weak diffuse scattering of the asymmetrical type. These two grounds lead to the expectations that in full-overlap rigor the diffuse scattering in the insect muscle would be circularly symmetrical and stronger than in the frog muscle. This is in line with a recent observation by K. C. Holmes (personal communication) that the diffuse scattering in the insect rigor muscle is rather strong and circularly symmetrical.

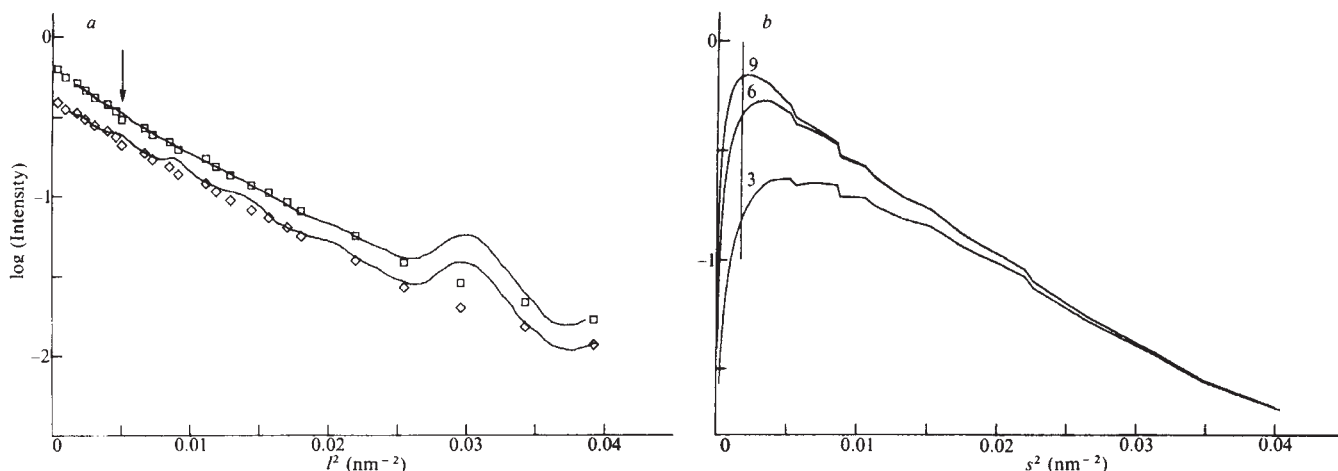


Fig. 4 *a*, Comparisons of model-calculated and observed diffuse disk intensity. *b*, Curves from model calculations showing the effect of varying the r.m.s. displacement (X) on the intensity distribution of the diffuse scattering. All data are shown as excentric Guinier plots³², $\log_{10}(I)$ against l^2 at $R = 0.0534 \text{ nm}^{-1}$. According to the Guinier approximation for a monodisperse system of low concentration, $I(s) = \exp(\frac{1}{3}(-4\pi s R_g)^2)$, where R_g is the radius of gyration of the particles, a straight line region should be obtained at small angles. For S-1 this applies out to a scattering angle of at least 12 mrad (ref. 33) corresponding to l^2 equal to about 0.005 nm^{-2} as indicated by the arrow. The observed data (—) are from patterns shown in Fig. 1c and d, the lower curve from the living relaxed and the upper from the non-overlap rigor muscle. The method of calculating the model data ($\diamond$) and ($\square$) is described below. It is seen that on going from the relaxed to the non-overlap rigor state the plot shifts upward with no change in its slope. The shift corresponds to an increase in the disk intensity by a factor of 1.5. The slope of the straight line will be the same in an excentric Guinier plot as in one which passes through the centre of the pattern and it provides a measure of R_g . To compare our results with the solution scattering data for S-1 we have calculated R_g using a more accurate procedure. We have also calculated the molecular volume (V) and the length of the mean chord (L). The formulae used were³⁴: $R_g^2 = \int_0^\infty p(r)r^2 dr / 2 \int_0^\infty p(r) dr$, $V = I(0)/4\pi \int_0^\infty s^2 I(s) ds$, and L was obtained by averaging the distance distribution function $p(r)$ where it is significantly above zero ($0 < r < 12 \text{ nm}$). $p(r) = 2 \int_0^\infty I(s)sr \sin(2\pi sr) ds$ was calculated after extrapolation of the intensity distribution $I(s)$ to $s = 0$ according to Guinier³² and to $s = 0.6 \text{ nm}^{-1}$ according to Porod^{36,37}. Taking the case of non-overlap rigor muscle where, in the absence of myosin layer lines, the treatment of the data is most straightforward, the values we obtain are: $R_g = 3.1 \text{ nm}$, $V = 150 \text{ nm}^3$ and $L = 4.0 \text{ nm}$ and these agree well with corresponding values obtained from solution studies^{16,17}, namely 3.28 nm , 151 nm^3 and 4.25 nm . To explain how the intensity of the diffuse scattering can change by a large factor without a shape change, we suggest a model where the displacements of myosin heads from their ideal lattice positions are defined by a linear combination of a narrow and a broad distribution: $P(x) = cP_1(x) + (1-c)P_2(x)$, where x is the displacement and c is the fraction of myosin heads with the narrow distribution $P_1(x)$. For our model we chose the simplest case, which has one infinitely narrow distribution ($P_1(x) = 1$ for $x = 0$ and 0 for $x > 0$) and one broad gaussian distribution with isotropic displacements ($P_2(x) = \exp(-2\pi(x/X)^2)$ where X is the r.m.s. displacement. Furthermore, it is assumed that the orientational and displacement disorder are independent of each other. This can be justified on the argument that a dependent relationship which changes the diffuse scattering significantly would also produce satellite peaks between the layer lines—an effect not seen in the muscle patterns. In what follows $\langle f^2(S) \rangle$ and $\langle f(S) \rangle$ mean, respectively, the average over all unit cells and the spherical average of the Fourier transform $f(S)$ of the myosin head. The intensity formulae are taken for Guinier³². Because the myofibrils are randomly oriented around the fibre axis, the intensity distributions obtained from the following equations should be cylindrically averaged about that axis. The general expression for the intensity of the diffuse scattering is:

$$I(S) = \overline{f^2(S)} - \overline{f(S)}^2 \quad (1)$$

We assume that at least the myosin heads with large X are randomly oriented. If the heads with small X have no orientational disorder equation (1) becomes:

$$I(S) = cf^2(S) + (1-c)\langle f^2(S) \rangle - [cf(S) + (1-c)\langle f(S) \rangle D(s)]^2 \quad (2)$$

Here $D(s) = \exp(-\frac{2}{3}(\pi s X)^2)$ is the Debye-Waller factor^{35,36}. This intensity distribution would not be circularly symmetrical on the film pattern unless c was very small and that turns out not to be the case. We therefore suggest that also the myosin heads with small X are randomly oriented. This is in agreement with the results of Thomas *et al.*²¹. The intensity equation then reduces to:

$$I(s) = \langle f^2(S) \rangle - [c\langle f(S) \rangle + (1-c)\langle f(S) \rangle D(s)]^2 \quad (3)$$

The intensity distributions for our model were calculated from equation (3) with $X = 9 \text{ nm}$ and using $c = 0$ ($\square$) and $c = 0.55$ ($\diamond$) respectively, for the non-overlap rigor and the relaxed states. As justified below, $\langle f(S) \rangle^2$ was approximated by $\langle f^2(S) \rangle$ which was taken from ref. 16 (and transposed as explained in Fig. 2 legend). With $X = 9 \text{ nm}$, $D(s)$ rapidly becomes very small with increasing s and this explains why our data then approach the solution scattering curve $\langle f^2(S) \rangle$. The intensity then approximates:

$$I(s) = (1-c^2)\langle f^2(s) \rangle \quad (4)$$

and on altering c the intensity change shows up as a parallel shift of the plot. It is justifiable to approximate $\langle f(S) \rangle^2$ by $\langle f^2(S) \rangle$ for small and intermediate values of s . If the molecule is spherical $\langle f(S) \rangle^2$ would be equal to $\langle f^2(S) \rangle$ at all values of s but with increasing elongation of the molecule a difference will occur at decreasing values of s . Model calculations using any reasonable dimensions of the myosin head demonstrate that the difference between the two terms is small over most of the region of the data shown. However, with increasing values of s , $\langle f(S) \rangle^2$ will eventually become very small compared with $\langle f^2(S) \rangle$ which means that the intensity will become equal to $\langle f^2(S) \rangle$ regardless of the value of c . It also means that rotational disorder alone will produce diffuse scattering, but this would be of very low intensity in the region where a high intensity is in fact observed. In *b*, the intensity distribution of the diffuse scattering was obtained from equation (3) given in *a*, using $c = 0$ and $X = 3, 6$ and 9 nm as indicated on the curves. The results are presented as Guinier plots. Note that with decreasing X the intensity drops below the continuation of the straight line region progressively away from the centre of the pattern, causing a strong curvature in the plot. Our data are measured as far inward as $s = 0.04 \text{ nm}^{-1}$ (vertical line) and no drop in intensity is seen. It is clear that this state of affairs is reproduced in the model when disordered myosin heads have a r.m.s. displacement of minimally 9 nm .

between a 'sticky patch' of the neck region of the myosin heads and light meromyosin (ref. 24 and J. Kendrick-Jones, personal communication). The myosin heads could then be either stuck to the backbone with movements restricted to rotation or detached with movements restricted only by the hinges near either end of the S-2 rod component. A similar effect could be due to an electrostatic interaction between the S-2 rod and the light meromyosin backbone. A third possibility would be a phase transition in the S-2 rod component between a rigid helix and a flexible coil²⁵. However, regardless of the physical explanation, the X-ray results indicate that ATP is necessary for the existence of the ordered population.

In full-overlap rigor we find that about 30% of the non-overlap rigor scattering in the meridional direction still remains (Fig. 3a, b). This scattering could be from two sources: (1) unattached heads with large displacement disorder and (2) some kind of disorder of the attached heads. Because the latter are firmly attached with a constant orientation¹⁹⁻²¹ they cannot have displacement or rotational disorder. The obvious remaining possibility is substitution disorder, which seems very likely, first, because actin monomers outnumber myosin heads^{4,8} and second, because in frog muscle attachment of myosin heads to actin can take place from three different myosin filaments. This allows a random mixture of actin monomers with and without myosin heads to occur, thus explaining the presence of diffuse scattering from attached myosin heads. In contrast, in insect flight muscle the myosin heads can attach from only two different myosin filaments and this imposes a strict marking scheme with a fairly high degree of order²⁶. For the frog muscle we calculate that the observed 30% intensity level of the non-overlap rigor scattering is consistent with a practically 100% attachment to actin (Fig. 3). This is in line with results using EPR^{19,21} and biochemical²⁷ techniques.

As regards structural information about myosin head attachment, our finding that a preferred orientation can be detected in the diffuse scattering will be relevant both for experiments which attempt to stop the cross-bridge cycle in a particular phase (for example, by using ATP analogues), and for studies of living contracting muscles. The advantage of analysing the diffuse scattering rather than the intensity distribution of the actin layer lines is that the latter requires a simultaneous solution of the structure of the actin filament, whereas the diffuse scattering depends only on the myosin head. The advantage will be greatest in the case of contracting muscle where the actin layer lines are unlikely to be much influenced by the

attachment of myosin heads because, due to asynchronous activity⁸, only a few will be attached at any given time. This effect will not restrict the study of the possible changes in the diffuse scatter. If during any step of the cross-bridge cycle enough myosin heads assume a preferred orientation this will be detectable from the intensity distribution of the diffuse scattering. In that respect, Figs 2b and 3b show what happens in full-overlap rigor. There are two instances where mention is made of diffuse scattering during contraction of frog muscle. The first¹¹ describes a slight increase of intensity in the region off the meridian where we observe a decrease in full-overlap rigor; the second²⁸ reports no changes. On our interpretation these results indicate that the rigor configuration of myosin heads does not constitute any significant part of the cross-bridge cycle. However, a final conclusion must await more detailed measurements of the shape of the diffuse scattering during contraction.

The availability of synchrotron radiation as a very intense X-ray source has made it possible to relate in time-resolved experiments mechanical activity during contraction to changes in the weaker parts of the low-angle pattern. Until recently the only parts of the pattern thus investigated were the stronger discrete reflections seen predominantly in striated muscles¹¹. Our finding that myosin heads produce diffuse scattering shows that, using synchrotron radiation, it will also be useful to study this part of the pattern during contraction and to do so even in muscles where the discrete reflections (particularly from myosin) are rather weak²⁹, or absent³⁰. The latter is the case in the living relaxed *Mytilus* anterior byssus retractor muscle, and using this muscle in time-resolved experiments we have already shown that changes in disk scattering during contraction provide important information about the relationship of mechanical activity to the behaviour of disordered myosin heads³⁰. Future applications of this approach might conceivably be found in experiments with muscle cells in tissue culture and with non-muscular contractile systems like slime moulds.

We thank Drs R. H. Abbott, O. Glatter, V. Luzzati, A. Miller and J. S. Wray for useful discussions, Ms R. S. Rapley for technical assistance, and Dr D. A. Blackburn, Professor G. F. Elliott and other members of staff at the Oxford Research Unit for their support. We are grateful for grants from the MRC, and the Muscular Dystrophy Association of the USA and to Aarhus University, Denmark, for the loan of X-ray equipment and for the provision of a research fellowship for F.R.P.

Received 13 September 1982; accepted 11 February 1983.

- Hanson, J. & Huxley, H. E. *Nature* **172**, 530-532 (1953).
- Huxley, A. F. & Niedergerke, R. *Nature* **173**, 971-973 (1954).
- Hanson, J. & Huxley, H. E. *Symp. Soc. exp. Biol.* **9**, 228-264 (1955).
- Huxley, H. E. *J. Biophys. Biochem. Cytol.* **3**, 631-648 (1957).
- Huxley, H. E. & Hanson, J. *Proc. Stockholm Conf. Electron Microscopy*, 202-204 (Academic, New York, 1957).
- Lowy, J. & Hanson, J. *Physiol. Rev.* **42**, 34-47 (1962).
- Lowy, S. in *Fibrous Proteins: Scientific, Industrial & Medical Aspects* (eds Parry, D. A. D. & Creamer, L. K.) 1-25 (Academic, New York, 1979).
- Huxley, H. E. & Brown, W. J. *molec. Biol.* **30**, 384-434 (1967).
- Squire, J. M. A. *Rev. Biophys. Bioengng* **4**, 137-163 (1975).
- Luther, P. K. & Squire, J. M. *J. molec. Biol.* **141**, 409-439 (1980).
- Huxley, H. E., Faruqi, A. R., Borda, J., Koch, M. H. J. & Milch, J. R. *Nature* **284**, 140-143 (1980).
- Huxley, H. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 2297-2301 (1981).
- Huxley, H. E., Faruqi, A. R., Kress, M., Borda, J. & Koch, M. H. J. *J. molec. Biol.* **158**, 637-684 (1982).
- Vibert, P. J., Haselgrove, J. C., Lowy, J. & Poulsen, F. R. *J. molec. Biol.* **71**, 757-767 (1972).
- Haselgrove, J. C. *J. molec. Biol.* **92**, 113-143 (1975).
- Mendelson, R. A. & Kretschmar, K. M. *Biochemistry* **19**, 4103-4108 (1980).
- Mendelson, R. A. *Nature* **298**, 665-667 (1982).
- Reedy, M. K., Holmes, K. C. & Tregear, R. T. *Nature* **207**, 1276-1280 (1965).
- Thomas, D. D., Ishiwata, S., Seidel, J. C. & Gergely, J. *Biophys. J.* **32**, 873-889 (1980).
- Borejdo, J., Assulin, O., Ando, T. & Putnam, S. J. *molec. Biol.* **158**, 391-414 (1982).
- Thomas, D. D. & Cooke, R. *Biophys. J.* **32**, 891-906 (1980).
- Scopes, R. K. in *Physiology and Biochemistry of Muscle as a Food* (eds Briskley, E. J., Cassens, R. G. & Trautmann, J. C.) 471-492 (University of Wisconsin Press, Madison, 1966).
- Davson, H. & Eggleton, M. G. in *Principles of Human Physiology*, 799 (Lea & Febiger, Philadelphia, 1962).
- Kendrick-Jones, J. *et al. J. Muscle Res. Cell Motility* **3**, 477 (1982).
- Sutoh, K., Sutoh, K., Karr, T. & Harrington, W. F. *J. molec. Biol.* **126**, 1-22 (1978).
- Holmes, K. C., Tregear, R. T. & Barrington Leigh, J. *Proc. R. Soc.* **207**, 13-33 (1980).
- Lovell, S. J., Knight, P. J. & Harrington, W. F., *Nature* **293**, 664-666 (1981).
- Yagi, N., O'Brien, E. J. & Matsubara, I. *Biophys. J.* **33**, 121-138 (1981).
- Matsubara, I. & Millman, B. M. *J. molec. Biol.* **82**, 527-536 (1974).
- Lowy, J. & Poulsen, F. R. *Nature* **299**, 308-312 (1982).
- Jewell, B. R. & Wilkie, D. R. *J. Physiol., Lond.* **152**, 30-47 (1960).
- Guinier, A. *X-ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies*, 185-294, 319-350 (Freeman, San Francisco, 1963).
- Kretschmar, K. M., Mendelson, R. A. & Morales, M. F. *Biochemistry* **17**, 2314-2318 (1978).
- Glatter, O. in *Small Angle X-ray Scattering* (eds Glatter, O. & Kratky, O.) 119-165 (Academic, London, 1982).
- Porod, G. *Kolloidzeitschrift* **124**, 51-57 (1951), **125**, 108-122 (1952).
- Debye, P. *Ann. Phys.* **43**, 49-95 (1914).
- Waller, I. *Z. Phys.* **17**, 398-408 (1923).

Intergalactic shock waves

V. V. Chernomordik & L. M. Ozerov

Department of Theoretical Physics, Lebedev Physical Institute,
Academy of Sciences, Moscow 117924, USSR

Young galaxies as well as quasars and active galactic nuclei, have long been considered as possible sources of shock waves entering the intergalactic medium¹⁻⁴. Following convincing evidence⁵ that the narrow absorption lines seen shortward of the Ly α emission line in very distant quasars are mostly intervening (and not intrinsic) Ly α absorption lines, various attempts⁶⁻⁸ have been made to relate the cold H I absorbers (in part, at least) with actual shock waves in the intergalactic medium. Here, we consider intergalactic shocks in a low density universe when energy losses from shocks are negligible at the redshifts of interest. However, even at this essentially adiabatic evolution stage, the absorption of the outer radiation in the outermost densest and coldest regions of the shock waves is capable of producing measurable Ly α absorption lines. Moreover, these lines should be observed, due to a specific geometry, in the form of special 'doublets'. We demonstrate that there are such doublets in quasar spectra among the Ly α 'forest', and this can apparently be regarded as the first conclusive evidence in favour of intergalactic shock waves.

Intergalactic gas filling the space between clusters of galaxies⁹ is thought to be highly ionized because the spectra show that quasars with redshifts $z \sim 2-3$ have no appreciable depression shortward of the Ly α emission line¹⁰. Current evidence strongly suggests that the ionization of the intergalactic gas is conditioned by the UV radiation from quasars so that the gas temperature is expected to be moderate: $T_g \approx 3 \times 10^4$ K (refs 11, 12) or even less in a low-density universe.

How does the gas density distribution change from its assumed initially homogeneous state when spreading a spherical shock wave through the intergalactic gas? We shall restrict ourselves here to the case when the metagalactic gas has a low density $\rho \leq (0.1-0.2)\rho_{cr}$, $\rho_{cr} \approx 10^{-29}$ g cm⁻³ being the critical density of the Friedman cosmological models. In this case, the energy losses of shock waves due to radiation and inverse Compton effect are insignificant at the redshifts of interest^{2,3}.

Even without radiation losses, noticeable density gradients are expected in shock waves. As is well known, in a strong shock wave the total gas density decreases rapidly from its maximum value at the front of the wave radius, while the gas temperature, in contrast, increases towards the epicentre of the explosion^{8,13}. Therefore, both the electron and neutral hydrogen concentrations, $n_e(r)$ and $n_{HI}(r)$, have their maximum near the wave front and decrease sharply to the explosion epicentre. The value of n_{HI} can be readily estimated if we take into account that (1) the time to approach ionization equilibrium (which at the redshifts under consideration is due to the rate of photoionization) is short compared with the recombination time (see ref. 5), and (2) the role of collisional ionization is negligibly small compared with the photoionization because in the conditions under consideration the temperature and density behind the wave front do not exceed $\sim 10^5$ K and 10^{-4} cm⁻³, respectively. As a result we get the relationship $n_{HI} \propto n_e^2 T^{-0.75}$ as used in ref. 5.

The fact that n_{HI} is expected to have a sharp maximum near the radius of the shock wave front should produce two identical Ly α absorption lines (or 'doublet') in the spectra of those quasars whose radiation reaches an observer through a shock wave zone (see Fig. 1). Let us estimate the expected doublet parameters for the adiabatic stage of the shock wave evolution.

The equivalent width of each component entering the Ly α doublet in the cosmologically co-moving, or rest frame related to the explosion epicentre, W_0 , can be easily estimated using

a convenient approximation³ for the shock wave radius $R_s(t)$ in the expanding intergalactic medium:

$$W_0 = A n_{HI} \delta R_s(t) / 3 \cos \theta$$

$$= \begin{cases} 0.13 \text{Å} \left(\frac{E_i}{10^{59} \text{erg}} \right)^{1/5} \left(\frac{\Omega_g}{0.1} \right)^{-1/5} h^{1/5} \left(\frac{1+z}{3} \right)^{0.3} \\ \quad \left(1 - \frac{t_e}{t} \right)^{2/5} \delta \frac{\tau_{HI}}{0.15} \left(\frac{\cos \theta}{\cos 70^\circ} \right)^{-1} & \text{at } \Omega = 1 \\ 0.11 \text{Å} \left(\frac{E_i}{10^{59} \text{erg}} \right)^{1/5} \left(\frac{\Omega_g}{0.1} \right)^{-1/5} h^{1/5} \\ \quad \left(1 - \frac{t_e}{t} \right)^{2/5} \delta \frac{\tau_{HI}}{0.15} \left(\frac{\cos \theta}{\cos 70^\circ} \right)^{-1} & \text{at } \Omega \ll z^{-1} \end{cases} \quad (1)$$

Here $A = 5.43 \times 10^{-15}$ Å cm² is determined by the oscillator strength for the Ly α line; n_{HI} is the mean H I density in the non-disturbed intergalactic gas; E_i is the initial explosional energy released at the moment t_e of an explosion; z is the cosmological redshift of a shock wave; θ is the angle between the line of sight and the normal to the wave front (see Fig. 1); δ is a structural coefficient depending on θ and the gas temperature behind the front, $T(R_s)$. Note that n_{HI} is as small as a few times 10^{-7} of the total gas density, $n_g = \Omega_g n_{cr}$ ($\Omega_g \sim 0.1$), at, for example, $z = 2$ when the optical depth of metagalactic H I in the Ly α line $\tau_{HI} = 0.1$. An exact relation between n_{HI} and n_g used in equation (1) depends also on cosmological parameters $\Omega = n_{total}/n_{cr}$ and $h = H/75$ km s⁻¹ Mpc⁻¹. As shown in ref. 14, for a wide set of the shock wave parameters $T(R_s)/T_g \leq 7$ so that $\delta \approx 0.5-1$ at $0 \leq \theta \leq \theta_{max} \sim 65-75^\circ$, it decreases sharply at $\theta > \theta_{max}$ and vanishes at $\theta = 90^\circ$.

The expected distance, $\Delta\lambda$, between two lines entering into a doublet is determined by the difference of radial velocities in the nearer and more distant (relative to an observer) parts of the shock front. As a result, the differential Doppler shift yields

$$\Delta\lambda = 2(v/c)\lambda \cos \theta = 2B \frac{R_s(t)}{ct} \cos \theta \quad (2)$$

v being the modulus of the gas velocity (near the wave front) relative to the explosion epicentre, and $B \approx 1$ is a numerical coefficient which depends only weakly¹⁴ on both the contraction of gas on the shock front and time t at $t \geq 2t_e$. Again, in the

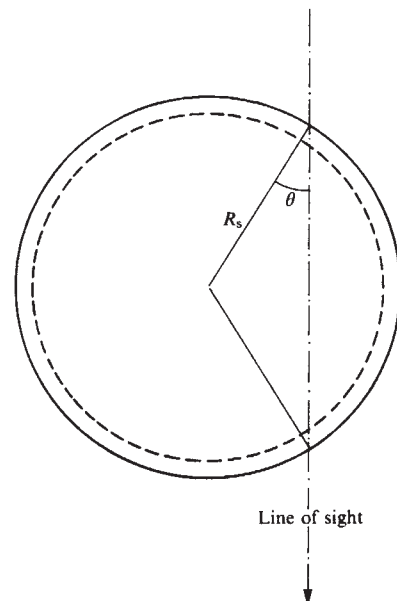


Fig. 1 Schematic outline of the shock front explaining the formation of an absorption doublet. The internal boundary of the most dense part of the shock wave zone is shown as a dashed line.

Table 1 Ly α doublets revealed in the available QSO samples

QSO	$\lambda_{\min}$ (Å)	$\lambda_{\max} = \lambda_{\text{Ly}\alpha}^{\text{emiss}}$ (Å)	$W_0^{\min}$ (Å)	Nos of the lines forming the Ly α doublets*	No. of lines forming into doublets	Total no. of lines in a sample with $W_0^{\min} \leq W_0 \leq 0.30$ Å
PKS0237-233 ($z_{\text{em}} = 2.223$)	3,730	3,918	0.09	5 & 6, 7 & 8, 11 & 13, 17 & 18, 21 & 22, 26 & 27	12	20
B21225+31.7 ($z_{\text{em}} = 2.200$)	3,400	3,890	0.16	19 & 20, 29 & 30, 35 & 36, 41 & 42, 54 & 55, 57 & 58	12	21
PKS2126-158 ($z_{\text{em}} = 3.280$)	4,380	5,200	0.22	67 & 68, 78 & 79, 99 & 100	6	11
PHL957 ($z_{\text{em}} = 2.690$)	3,650	4,486	0.16	64 & 65, 80 & 82	4 [†]	10
QSO0421+019 ($z_{\text{em}} = 2.051$)	3,370	3,709	0.16	12 & 13, 18 & 20, 21 & 22, 31 & 33	8	12
QSO0002+051 ($z_{\text{em}} = 1.899$)	3,360	3,524	0.16	9 & 10	2	4
QSO0119-046 ($z_{\text{em}} = 1.937$)	3,280	3,570	0.16		0 [‡]	7

* As given in refs 5 and 16.

† Identification of heavy elements (C, Si, and so on) lines in the PHL957 spectrum which, according to ref. 5, is still uncertain so that the true number of doublets may prove to be even higher¹⁴ (nos 58 & 59, 61 & 63, 65 & 66).

‡ One doublet (nos 1 & 2 as given in ref. 17) appears in our sample if we take a slightly smaller $W_0^{\min} = 0.15$ Å. Within closer ranges $3,410 \leq \lambda \leq 3,570$ Å a better resolution enables one to analyse narrower absorption lines. As a result, one finds there two doublets (nos 15 & 17, 26 & 27) so that four lines of seven with $0.09 \leq W_0 \leq 0.16$ Å form into doublets.

rest frame related to the explosion epicentre, we have from equations (1) and (2)

$$\Delta\lambda_0 = (1+z)^{-1} \Delta\lambda = KW_0 \cos^2 \theta, \quad (3)$$

where $K \approx 6B/\delta\tau_{\text{HI}}$ if $\Omega \ll z^{-1}$, and $K \approx 9B/\delta\tau_{\text{HI}}$ if $\Omega = 1$, $\tau_{\text{HI}} \sim 0.1$ being the optical depth of metagalactic H I in the Ly α line. An average value of $\cos^2 \theta$ for a large sample of shock waves located on the line of sight can be easily estimated to be

$$\langle \cos^2 \theta \rangle = \cos^4 \theta \big|_{\theta_{\min}}^{\theta_{\max}} / \cos 2\theta \big|_{\theta_{\min}}^{\theta_{\max}} \quad (4)$$

where $\theta_{\min} = \theta_{\min}(E_i)$ is determined by comparison of W_0 calculated from equation (1) with the equivalent width, $W_{\min}$, of the most narrow lines which could be reliably registered. Practically, $W_{\min} \sim 0.1$ Å.

Therefore, cosmological shock waves created by explosions in young galaxies and quasars are expected to produce absorption lines with the following features: (1) a significant part of sharp absorption lines should form very close Ly α doublets ($\Delta\lambda_0 \leq 3$ Å), and (2) both components of a doublet should have the same equivalent width within the range $0.01 \leq W_0 \leq 0.3$ Å.

Are there any such peculiar doublets in the absorption spectra of distant quasars?

Observations⁵ of several distant quasars revealed in their absorption spectra the so-called Ly α 'forest' whose origin is thought to be due to absorptions in the intergalactic gas at redshifts $z \sim 2-3$ (refs 5, 15). Practically all discovered Ly α lines have equivalent rest widths (in the reference frame related to an absorbing material) within the range $W_0 = 0.06-1$ Å, and the line density $n(W)$ (per unit Δz per unit W) is decreasing rapidly with W . Of course, the true line density $n^*(W) = (N^*/W^*) \exp(-W/W^*)$ should differ from the observed one, $n(W)$, owing to blending⁵. The blends formed from more numerous lines of a smaller equivalent width make up within a sample of absorption lines with $W_1 \leq W_0 \leq W_2$ a fraction equal to¹⁴:

$$[p(x_1)\tilde{N}(W_1) - p(x_2)\tilde{N}(W_2)] / [\tilde{N}(W_1) - \tilde{N}(W_2)]$$

where $\tilde{N}(W_0) = \int_{W_0}^{\infty} n(W_i) dW_i$ is the true (corrected for blending) number of absorption lines with $W_i \geq W_0$, and $p(x) = a(x^2 + 2xe^{-x} + 2)/[1 + a(x^2 - 2)]$, $x \equiv W_0/W_0^*$, $a = 3 \times 10^{-3} N^*W^*$. For the sample presented in ref. 5 ($N^* = 224$ and $W^* = 0.22$ Å) we have $W_1 = 0.09$ Å and $W_2 = 0.3$ Å so that the expected percentage of blends is 53%. As for yet unidentified, wider lines ($W_0 > 0.3$ Å), they might be mainly blends.

To elucidate the physical nature of the absorption lines with $\lambda < \lambda_{\text{Ly}\alpha}^{\text{emiss}}$, observations are needed with a spectral resolution good enough to analyse very narrow lines with $W_0 < 0.3$ Å. In the sample of six quasars investigated in ref. 5, the necessary observational information is available only for PKS0237-233, B21225+31.7, PKS2126-158, and, possibly, for PHL957.

The doublets having the expected parameters in the spectra of these quasars and of three other quasars investigated in a more recent work^{16,17} are listed in Table 1. The results show that in the available quasar spectra of high quality, about half of all narrow absorption lines with $\lambda < \lambda_{\text{Ly}\alpha}^{\text{emiss}}$ and $W_0 \leq 0.3$ Å form doublets consisting of components of an equal width with a rather small distance, $\Delta\lambda$, between them (as a rule these components appear as neighbouring lines in the spectrum).

Note that the fact most (half or so) all narrow lines enters into doublets is evidence for statistical significance of most of these doublets. For example, in the PKS0237-233 spectrum obtained with the best resolution (and containing, as a result, the largest number of narrow lines) the expected fraction of neighbouring lines fortuitously having the same W should not exceed 2/7 (assuming that they have uniform distribution in W between 0.3 and 0.9 Å) while the observed fraction is 3/5, that is twice as much. Moreover, the observed fraction of doublets consisting of the broadest lines of the same $W = 0.7-0.9$ Å is even greater (2/3), while the fraction of such chance doublets is expected to decrease with W due to the rareness of the occurrence of broad lines. Of course, a more sophisticated statistical analysis of what fraction of doublets is real in each quasar spectrum is highly desirable, but it requires a much greater sample size.

Figure 2 demonstrates a relationship between W_0 and $\Delta\lambda_0$ for the doublets found. Although the statistics are still rather poor, the available data do not contradict the expected relationship $\Delta\lambda_0 \approx 8W_0$ shown by the dashed line. The latter is obtained by the averaging of equation (3) at $\tau_{\text{HI}} = 0.15$, $\delta = 0.75$, $E_i = 4 \times 10^{59}$ erg, and $W_{\min} = 0.09$ Å. The scattering of observational data around the dashed line may be explained by the various θ for the individual shocks.

Therefore, we believe that the Ly α doublets discovered in the spectra of the quasars listed in Table 1 are, indeed, conditioned by the shell structure of intervening absorbers (although the sample size on which this conclusion is based is still not too large). These shells seem to be related to the intergalactic shock waves. The rest of the narrow Ly α lines that have not formed into doublets can be explained essentially by a sum of blends (whose expected percentage has been estimated above) plus the Ly α lines produced by the assumed intergalactic H I clouds⁵.

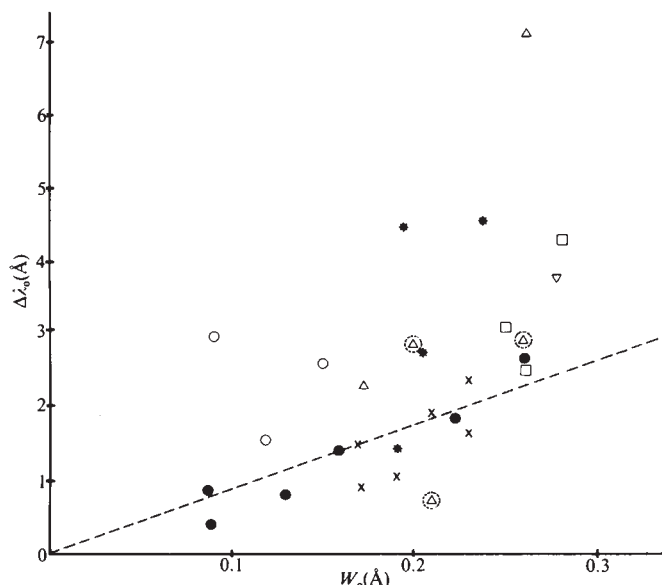


Fig. 2 A plot of distance at rest between the components of a doublet against the equivalent width at rest. ●, PKS0237-233; ×, B21225+31.7; △, PHL957; △, correspond to those pairs of absorption lines in PHL957 whose doublet interpretation is impossible if the preliminary identification⁵ of these lines as being produced by heavy elements were confirmed; ★, QSO0421+019; □, PKS2126-158; ▽, QSO0002+051; ○, QSO0119-046.

By using the mean radial distance between the Ly α absorption lines found in ref. 5 we can easily estimate the spatial density, n_s , of the progenitors for the shocks. Reducing this to the present, it appears that $n_s|_{z=0} \sim 10^{-75} \text{ cm}^{-3}$. Such coincidence by an order of magnitude with the spatial density of normal galaxies seems to agree with the hypothesis advanced earlier (see refs 2, 3) that the explosions in young galaxies (which took place presumably at the end of the 'bright phase', at $z \sim 10$) led to the formation of the intergalactic two-phase structure consisting of a large number of spherical cavities (hot rarefied regions) surrounded by rather thin, but dense and cold shells.

At the time corresponding to $z \approx 2$, these cavities should fill the main volume fraction, f , of the intergalactic medium, because $f = (4/3)\pi R_s^3 n_s|_{z=2} \sim 1$. By this time, most shocks have not intersected their neighbours. Therefore, a significant fraction of the absorption lines can be indeed observed in the form of doublets. In contrast, at smaller redshifts not only the doublet character of narrow absorption lines, but the very presence of numerous Ly α lines is hardly possible since the mutual intersections of shock waves are capable of disrupting any regular structure formed previously [It is tempting to speculate that this may explain decrease of the Ly α line density at $z < 2$ noted in ref. 16.]

To conclude, we stress the necessity of further searches for the absorption Ly α doublets among the most narrow absorption lines with $\lambda < \lambda_{\text{Ly}\alpha}^{\text{emiss}}$ in the spectra of very distant quasars ($z \geq 2$). If observations with the same, or still higher, resolution confirm the hypothesis advanced above that such doublets are conditioned by absorptions in intergalactic shock waves there would appear to be an opportunity of directly investigating the physical properties of intergalactic gas and its evolutionary history.

A referee has pointed out that an analysis of the Ly α lines produced in metagalactic shocks (without any direct observational evidence for the model such as presented above) has been carried out in an unpublished paper by S. Ikeuchi, K. Tomisaka and J. P. Ostriker.

Received 9 December 1982; accepted 22 February 1983.

- Schwarz, J., Ostriker, J. P. & Yahil, A. *Astrophys. J.* **202**, 1 (1975).
- Ozernoy, L. M. & Chernomordik, V. V. *Soviet. astr. Lett.* **2**, 145 (1976); (*Pis'ma Astr. Zh.* **2**, 375; 1976).

- Ozernoy, L. M. & Chernomordik, V. V. *Soviet. Astr.* **22**, 141 (1978); (*Astr. Zh.* **55**, 236; 1978).
- Ikeuchi, S. *Publ. astr. Soc. Jap.* **33**, 211 (1981).
- Sargent, W. L. W., Young, P. J., Boksenberg, A. & Tytler, D. *Astrophys. J. Suppl.* **42**, 4 (1980).
- Ostriker, J. P. & Cowie, L. L. *Astrophys. J. Lett.* **243**, L127 (1981).
- Shull, J. M. & Silk, J. *Astrophys. J.* **249**, 26 (1981).
- Ikeuchi, S. & Tomisaka, K. Preprint (Hokkaido University, 1982).
- Field, G. B. Preprint (Center for Astrophysics, No. 1132, 1978).
- Gunn, J. E. & Peterson, B. A. *Astrophys. J.* **142**, 1633 (1965).
- Sherman, R. D. *Astrophys. J.* **237**, 355; **240**, 737 (1980).
- Bergeron, J. & Salpeter, E. E. *Astrophys. Lett.* **7**, 115 (1970).
- Cox D. P. *Astrophys. J.* **178**, 159 (1972).
- Chernomordik V. V. & Ozernoy L. M. *Astrophys. Space Sci.* (in the press).
- Weymann, R. J., Carswell, R. F. & Smith, M. G. A. *Rev. Astr. Astrophys.* **19**, 41 (1981).
- Young, P. J., Sargent, W. L. W. & Boksenberg, A. *Astrophys. J.* **252**, 10 (1982).
- Sargent, W. L. W., Young, P. J. & Boksenberg, A. *Astrophys. J.* **252**, 54 (1982).

A new test of the cosmological interpretation of QSO redshifts

P. A. Shaver

European Southern Observatory, Karl-Schwarzschild-strasse 2,
D-8046 Garching bei München, FRG

J. G. Robertson

Anglo-Australian Observatory, PO Box 296, Epping,
New South Wales 2121, Australia

Close QSO pairs can be used to test the cosmological interpretation of QSO redshifts. If (some of) the narrow absorption lines in QSO spectra arise in intervening matter, and if QSOs are surrounded by gaseous haloes or located in regions of high matter density (for example, clusters of galaxies), one may expect to find absorption in the spectrum of one QSO at the redshift of a foreground QSO which happens to be located nearby on the plane of the sky ('associated absorption'). According to the cosmological interpretation, the foreground QSO should have the lower redshift. If, on the other hand, QSO redshifts are unrelated to their distances, one would expect to find as many cases in which the foreground QSO has the higher redshift of the two; or if the absorption lines are not even due to intervening matter, there should be no coincidences of absorption and emission redshifts beyond random expectation. Thus, by using associated absorption to distinguish which QSO of a pair is in front of the other, we have a simple and straightforward test of the cosmological interpretation.

The symmetry of QSO pairs—the fact that both members are QSOs—renders them particularly powerful tools for testing the cosmological interpretation of QSO redshifts. QSO–galaxy pairs, for example, do not share this symmetry, and can only be used to demonstrate that QSOs are at least as distant as the galaxies (in principle, the QSOs could be embedded in the haloes of the galaxies).

The QSO pairs which are presently available are listed in Table 1. In all cases the relevant absorption lines with equivalent widths of at least 1 Å could be detected over the appropriate redshift range. Cases where confusion is likely (for example, in the Ly α 'forest') have been excluded. Table 1 is inhomogeneous insofar as different lines (Mg II, C IV and Ly α) have been used in different cases, but we note that the equivalent width of Ly α is greater than that of C IV 1,548 by a factor of three on average, so an absence of Ly α absorption probably also implies an absence of C IV absorption.

Four cases have been found in which associated absorption is present at z_{em} (lower) (within $|z_{\text{abs}} - z_{\text{em}}|/(1 + z_{\text{em}}) = 0.0067$, which corresponds to 2,000 km s $^{-1}$ if the redshifts are Doppler, this being about the maximum for which correlations between galaxies might be expected; see refs 1, 2). Considering the average number density of heavy-element absorption systems², the probability of these coincidences occurring by chance with a random distribution of redshifts is $< 10^{-3}$. Furthermore, they occur mostly among the QSO pairs having the smallest angular

Table 1 Associated absorption in close pairs of QSOs

QSO pair	Separation (arc min)	Associated absorption at z_{em} (lower)			Associated absorption at z_{em} (higher)		
		z_{em}	Relevant lines	$\frac{ z_{abs} - z_{em} }{(1 + z_{em})}$	z_{em}	Relevant lines	$\frac{ z_{abs} - z_{em} }{(1 + z_{em})}$
1548+114A, B*	0.1	0.44	MgII	0.0047	1.90	—	—
2359-022A, B†	1.0	0.86	—	—	2.81	Ly α	>
0307-195A, B‡	1.0	2.12	C IV, Ly α	0.0003	2.14	C IV, Ly α	>
0254-334A, B§	1.0	1.86	C IV	>	1.91	C IV	>
0028,9+003	1.0	1.73	C IV	0.0006	2.22	C IV	>
2206-199A, B†	2.1	1.58	—	—	2.55	Ly α	>
1623+268,9¶	2.9	2.52	—	—	2.61	Ly α	>
1228+076,7#	3.4	1.88	C IV	0.0067	2.39	—	—
1604+176,7†	4.4	2.04	C IV	>	2.32	—	—
2224-408A, B†	4.8	1.95	C IV	>	2.33	Ly α	>
0203-497,8†	5.7	1.44	C IV	>	2.60	Ly β	>

> Indicates $|z_{abs} - z_{em}|/(1 + z_{em}) > 0.0067$; — indicates that the data are either inadequate or ambiguous, and these cases were not included in the statistical calculations.

* Ref. 3; † this study; ‡ ref. 4; § ref. 5; || ref. 6; ¶ ref. 7; # ref. 8.

separations. This provides strong evidence that these absorption lines are due to intervening matter, and that at least some QSOs are surrounded by diffuse matter (for example, extended haloes or clusters of galaxies).

On the other hand, no cases have been found of associated absorption at z_{em} (higher). There seems to be a clear asymmetry: associated absorption is sometimes found at z_{em} (lower) but never at z_{em} (higher). Application of the Fisher exact probability test (which uses the presence or absence of associated absorption on a simple yes-no basis) to Table 1 shows that the probability that QSO redshifts are randomly distributed with distance is only 0.04. It appears that the foreground QSO always has the lower redshift.

Indeed, the mere existence of the phenomenon of associated absorption presents severe obstacles (of an astrophysical nature) to non-cosmological interpretations of redshifts. The fact that several cases are now known firmly establishes that these are not chance coincidences—the absorption is really associated with the foreground QSO. A non-cosmological (emission) redshift of that QSO would therefore require an identical non-cosmological origin for the associated absorption redshift, yet the absorption presumably arises in a region of quite different physical conditions. Furthermore, if many absorption systems are intervening (the present evidence based on associated absorption may soon be augmented by searches for common absorption at the same redshift in both spectra of very close pairs of QSOs), the absence of absorption systems with far higher redshifts than the emission redshifts of individual QSOs will provide a very strong argument against noncosmological interpretations.

While the present results are not totally conclusive, it should be noted that the sample of QSO pairs can easily be expanded, making possible far stronger conclusions (extending also the sub-samples of QSOs) and other variations of these tests. These results, of course, say nothing about the actual relationship between redshift and distance, but they do indicate that higher-redshift QSOs are more distant than lower-redshift QSOs, and, if substantiated by further observations, this will be sufficient to eliminate the hypothesis that QSO redshifts are unrelated to distance.

Received 21 February; accepted 14 March 1983.

1. Sargent, W. L. W., Young, P. J., Boksenberg, A. & Tytler, D. *Astrophys. J. Suppl.* **42**, 41 (1980).
2. Young, P., Sargent, W. L. W. & Boksenberg, A. *Astrophys. J. Suppl.* **48**, 455 (1982).
3. Burbidge, E. M., Smith, H. E., Weymann, R. J. & Williams, R. E. *Astrophys. J.* **218**, 1 (1977).
4. Shaver, P. A. & Robertson, J. G. *Astrophys. J. Lett.* (in the press).
5. Wright, A. E., Morton, D. C., Peterson, B. A. & Jauncey, D. L. *Mon. Not. R. astr. Soc.* **199**, 81 (1982).
6. Shaver, P. A., Boksenberg, A. & Robertson, J. G. *Astrophys. J. Lett.* **261**, L7 (1982).
7. Sargent, W. L. W., Young, P. & Schneider, D. P. *Astrophys. J.* **256**, 374 (1982).
8. Robertson, J. G. & Shaver, P. A. Preprint.

Discovery of DHM0054-284, a quasar of redshift 3.61

T. Shanks, R. Fong & B. J. Boyle

Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

We report here the discovery of a quasar of redshift 3.61 detected in a broad band survey of stars using COSMOS measurements of UK Schmidt Telescope (UKST) photographs. The identification of the object as a quasar was confirmed by spectroscopic observations made with the Anglo-Australian Telescope (AAT). The quasar's redshift is the second highest so far detected and the highest for a quasar discovered using purely optical techniques. We describe here the search for DHM0054-284 and then suggest that the methods provide an efficient way of finding more high-redshift quasars.

The procedure began with the measurement by the COSMOS machine¹ of U , B , V and R UKST photographs to obtain magnitudes and positions for 20,000 stars with $B < 21$ mag in a $4^\circ \times 4^\circ$ area of sky. The area of sky surveyed was centred on the south galactic pole and the COSMOS magnitudes were calibrated using standard star sequences in the field^{2,3}. An objective prism photograph of this area of sky had previously been surveyed by eye in a search for emission line quasars⁴. This revealed the positions of 73 candidate quasars in our measured region. Many of these quasars have redshifts estimated from the objective prism spectra. By identifying these images we were able to find their colours in our magnitude system and these proved to be of great help in indicating the likely colours of higher redshift quasars.

Figure 1 plots $U-B$, $B-V$ and $V-R$ colour against redshift for every quasar identified in the prism survey that had an estimated redshift and that was identified by COSMOS as a star (44 images). The colours are zero pointed in the Johnson system (although the B photographic passband, in particular, is wider than the standard Johnson passband). The estimated r.m.s. error on the colours are ± 0.25 mag. Full details of the photometry techniques used will be given elsewhere. Although not all of the objective prism quasars are confirmed as such, where slit spectra have been obtained the proportion of quasars has been found to be $> 90\%$. In these cases, the slit spectra show that the objective prism estimates of the redshifts are accurate to ± 0.1 .

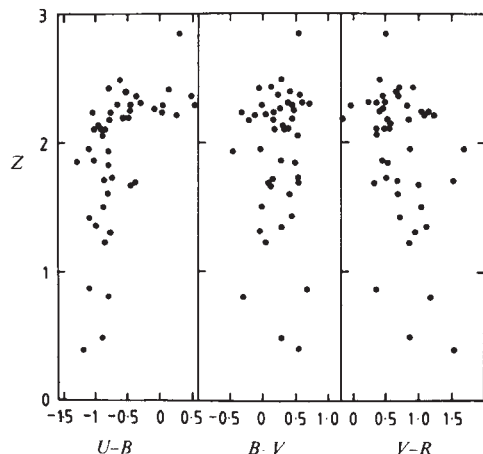


Fig. 1 Colour-redshift diagrams for the emission line quasars detected by R. G. Clowes and A. Savage on the South Galactic Pole Field.

Figure 1 shows that, in general, quasar colours stay reasonably constant as a function of redshift. For the $B-V$ and $V-R$ colours this seems to be true for the whole redshift range shown, with the average colours, $B-V = 0.2 \pm 0.3$ mag and $V-R = 0.7 \pm 0.4$ mag. Quasars with $z < 2.2$ have $U-B$ colours which also stay constant with redshift, with the average $U-B = -0.9 \pm 0.2$ mag. However, in Fig. 1 quasars with higher redshifts on average tend to have redder $U-B$ colours. There seems to be no straightforward selection effect which might have caused this result. At redshift 2.2, the $\text{Ly}\alpha$ line has been redshifted half way through the U band and thus the colour change may represent a relative dimming in this band caused by the decrease in continuum flux found shortwards of $\text{Ly}\alpha$ in many quasars. This effect has been noted previously⁵ and it causes the otherwise excellent UV excess method of quasar detection to be biased against the discovery of high-redshift quasars. The results in Fig. 1 suggest that the only change in broad-band colours with redshift occurs when the $\text{Ly}\alpha$ line is being shifted through the bluer passband. If so, then this implies that the average quasar $B-V$ colour will stay constant to $z \sim 2.9$ when the colour may redden by more than 1 mag. Similarly, the average $V-R$ colour may stay constant $z \sim 3.6$ before it also changes. Thus, to find quasars of redshift $z \sim 3.5$, the prescription would seem to be to look at stars with $V-R$ colour near the average of 0.7 mag and $B-V$ colour redder than 1.0 mag. It is now well established that at faint limits ($B \sim 20$ mag) stars are dominated by two populations, one a halo population with $B-V \sim 0.7$ mag and $V-R \sim 0.7$ mag and the other a disk population with $B-V \sim 1.3$ mag and $V-R \sim 1.1$ mag (refs 6,7). To leave a manageable number of candidate quasars with the least possible contamination by these galactic stars we confined our attention to stars with $B < 20$ mag, $B-V > 1.0$ and $V-R < 0.6$ mag. The prism spectra of about 75 of these stars were then inspected by eye. Most were distinctly red but otherwise featureless and these were taken to be galactic stars, spread into this region of the colour-colour diagram by r.m.s. errors in our magnitudes. Two candidates, however, had more unusual spectra, one of which was DHM0054-284 whose objective prism spectrum is shown in Fig. 2. The prism spectrum was noted as being very stubby with some suggestion of an emission line.

On the nights of 20-21 August 1982, low dispersion (156 Å per mm) spectra through a 2.0 arc s slit were obtained for the two prime candidate objects using the AAT with the Royal Greenwich Observatory spectrograph and Image Photon Counting System detector. The first candidate had the spectrum of a galactic star whereas DHM0054-284 had the spectrum shown in Fig. 3. The broad emission line has a wavelength of 5,603 Å and there is a sharp cutoff in the spectrum at 4,205 Å. These two features have wavelengths in the right ratio to be,

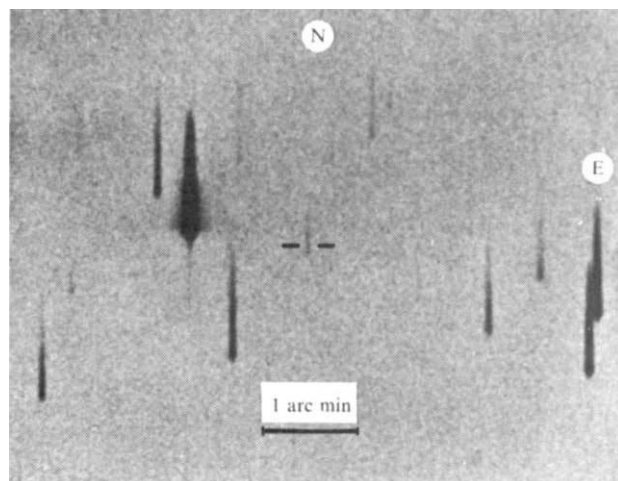


Fig. 2 Objective prism spectrum of DHM0054-284 taken from UKST plate UJ3682P. This photograph will also serve as a finding chart for the object whose equinox (1950) coordinates are RA 00 h 54 min 00 s, dec. $-28^{\circ}24'45''$.

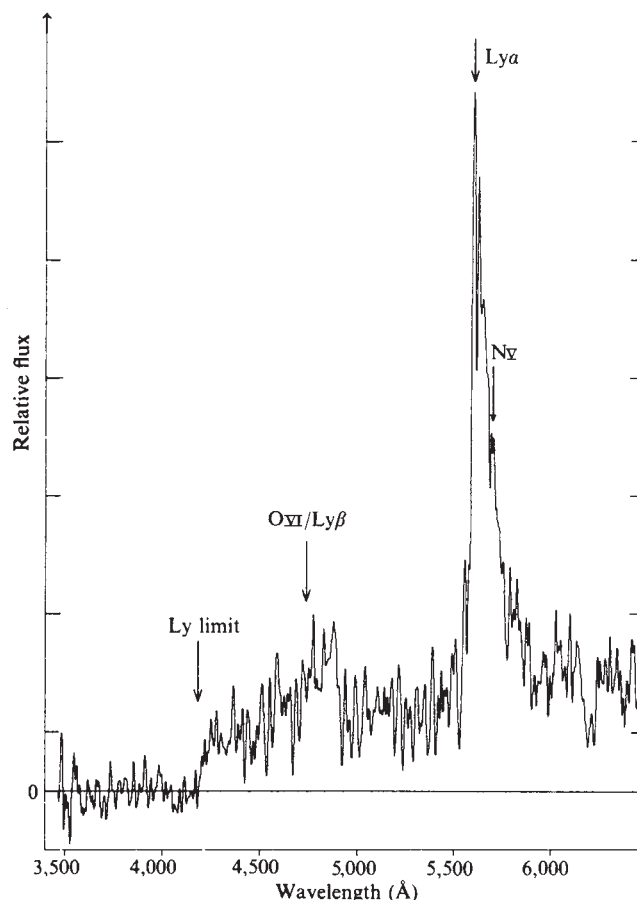


Fig. 3 Spectrum of DHM0054-284 obtained at 156 Å per mm on the AAT with the RGO spectrograph and the IPCS detector. Shortwards of the $\text{Ly}\alpha$ emission line, many absorption features can be seen in the spectrum. There is also an absorption feature at nearly the same wavelength as the $\text{Ly}\alpha$ emission line.

respectively, the $\text{Ly}\alpha$ (1,216 Å) emission line and the Lyman limit (912 Å) in a quasar of redshift $3.61 \pm .01$. The $\text{Ly}\alpha$ line lies past the red cutoff in the objective prism spectrum and the emission line seen there is a very broad feature in this spectrum, coinciding with the position of the $\text{OVI}/\text{Ly}\beta$ line at redshift 3.61. The Nv (1,240 Å) line may also be present in the spectrum at 5,701 Å. A quasar at redshift 3.61 would also be expected to show the CIV (1,550 Å) line at 7,141 Å, in a range

where our original observational set up had little sensitivity. Subsequently a red spectrum of this quasar was obtained by Drs Zuiderwijk and de Ruiter on the European Space Observatory 3.6-m telescope in which the C IV line is probably detected, at a low signal-to-noise ratio. We also observed eight other candidate objects that had similar colours to DHM0054-284 but less interesting prism spectra, but these all turned out to be red, galactic stars. The magnitude and colours of DHM0054-284 are $B = 19.55$, $B - V = 1.3$ and $V - R = 0.5$.

Thus, after a period of 10 yr when the highest quasar redshift remained at $z = 3.53$ (ref. 8), two quasars with higher redshifts have now been found in quick succession. The quasar with the highest redshift, PKS2000-330 with $z = 3.78$, was detected first as a radio source⁹, whereas DHM0054-284 has been found by purely optical methods. The difficulty there has been in detecting QSOs with redshift > 3.5 by any method indicates the efficiency of the techniques described here. We are continuing this work, looking for stars with red $V - R$ colours, in the hope of detecting even higher redshift quasars.

We thank R. G. Clowes and A. Savage for use of their quasar survey in advance of publication. The COSMOS and UK Schmidt telescope Units at The Royal Observatory, Edinburgh, are acknowledged for assistance. We thank E. Zuiderwijk and H. de Ruiter for making the additional observation on the ESO 3.6-m La Silla telescope. Finally T.S. and B.J.B. acknowledge SERC financial support during this work.

Received 18 February; accepted 14 March 1983.

1. Stobie, R. S., Smith, G. M., Lutz, R. K. & Martin, R. *Proc. int. Workshop on Image Processing in Astronomy*, Trieste (1978).
2. Cannon, R. D. *Mon. Not. R. astr. Soc.* **167**, 551-580 (1974).
3. Hawkins, M.R.S. *Mon. Not. R. astr. Soc.* **194**, 1013-1023 (1981).
4. Clowes, R. G. & Savage, A. *Mon. Not. R. astr. Soc.* (in the press).
5. Cheney, J. E. & Rowan-Robinson, M. *Mon. Not. R. astr. Soc.* **195**, 497-504 (1981).
6. Kron, R. G. thesis, Univ. California, Berkeley (1978).
7. Shanks, T., Philipps, S. & Fong, R. *Mon. Not. R. astr. Soc.* **191**, 47P-52P (1980).
8. Wampler, E. J., Robinson, L. B., Baldwin, J. A. & Burbidge, E. M. *Nature* **243**, 336-337 (1973).
9. Peterson, B. A., Savage, A., Jauncey, D. L. & Wright, A. E. *Astrophys. J. Lett.* **260**, L27-L29 (1982).

Optical polarization position angle versus radio structure axis in Seyfert galaxies

Robert R. J. Antonucci

National Radio Astronomy Observatory, Edgemont Road, Charlottesville, Virginia 22901, USA

Optical polarization position angles tend to align with large-scale radio structure in low polarization quasars¹, and they tend to align with or be perpendicular to the radio structure in radio galaxies². In some cases active object emission lines are polarized like the continuum, so the polarization is probably caused by scattering (ref. 3 and R.R.J.A., in preparation). Thus, the polarization alignments indicate the geometrical distribution of scatterers in those cases and so provide information on the matter distribution in the innermost regions. In a new interpretation of the polarization of the Seyfert 2 galaxy NGC1068 (ref. 4), it is noted that the optical polarization of that object is perpendicular to the inner radio contours, while the polarization of the Seyfert 1 galaxy NGC4151 is parallel to the associated radio axis. It is speculated that these are prototypical of two polarization classes of Seyferts analogous to the perpendicular and parallel radio galaxy groups. I show here that this is indeed the case: all Seyfert 1 galaxies in this sample have roughly parallel polarization and all Seyfert 2 galaxies have roughly perpendicular polarizations. These alignment effects can be interpreted as being due to thin and thick scattering disks, respectively, surrounding the continuum sources. This would represent a fundamental difference between the two types of Seyfert galaxies.

The optical polarimetry data used in this study and the determination of which objects qualify as Seyferts are taken from ref. 5. As low optical polarizations present problems of poor accuracy and contamination by galactic interstellar polarization, the following screening method was used. First, 33 of their 99 polarizations were discarded because the position angle errors $\sigma(\theta) > 10^\circ$. Next, 17 more were discarded because the polarization $P < 0.6\%$, so that it would be difficult to determine whether the data were contaminated by interstellar polarization. For the remaining 49 objects, the galactic coordinates and the polarization position angles with respect to the galactic pole were computed, then plotted on the interstellar-polarization map of Matthewson and Ford⁶. The 25 having very little interstellar polarization in the vicinity were considered to be intrinsically polarized, and included in the sample. The objects Markarian 3, NGC1275 and 3C390.3 have neighbouring stars that are rather polarized, but at different position angles than the galaxies themselves; because of spectropolarimetry³ and variability⁷ arguments, their polarizations are probably mostly intrinsic and they were therefore retained.

Next, I examined the available radio maps to determine whether an axis could be measured; this was possible in only 11 cases. Sources for the radio data are listed in Table 1. Error estimates for the structure position angles based on the map appearance are also given. Because there were usually only a few resolution elements across the source, substantial additional errors are possible with twisted sources. For example, earlier, lower resolution maps of NGC1068 would have given a rather different position angle.

Table 1 shows that Seyfert 1 optical polarizations tend to be parallel to the radio axes, while Seyfert 2 polarizations tend to be perpendicular. According to the Kolmogorov-Smirnov test, the two-tailed statistical significance level is $< 0.5\%$ for the Seyfert 1s and $< 2.0\%$ for the Seyfert 2s. If we exclude the rather problematical objects (see Table 1) the significance levels are the same.

We cannot be sure of the origin of the optical polarizations in these objects. Angel *et al.*⁸ concluded that in NGC1068 dust scattering is the cause, partially because of a strong rise in polarization to the blue, but that interpretation has been disputed⁴. Thompson *et al.*³ concluded that dust scattering causes the polarization of several other Seyferts, but perhaps some of these should be reconsidered in view of the new NGC1068 interpretation⁴, that the rise in P to the blue is due to wavelength-dependent dilution by unpolarized starlight. However, the polarization of the $H\alpha$ emission lines⁹ does favour some scattering process so electron scattering should be considered. From high-resolution spectropolarimetry of NGC4151, Schmidt and Miller⁷ found that the continuum is a combination of unpolarized starlight and a power law component with wavelength-independent polarization; this suggests synchrotron radiation or electron scattering. Finally, the very high, wavelength-independent optical polarization of the 'perpendicular group' radio galaxy 3C234 (R.R.J.A., in preparation), present in the $H\alpha$ line as well as the continuum, argues for electron scattering in that object.

Nor can we be sure of the cause of the radio emission, with both radio galaxy-like jets and supernovae proposed as explanations. The polarization alignments seem to favour the former explanation because they strengthen the similarity of Seyfert radio emission to that in radio galaxies and quasars.

As noted by Stockman *et al.*, optically thin disks scattering nuclear light would result in parallel polarization while optically and geometrically thick disks would result in perpendicular polarization. Alternatively, an optically thick but geometrically thin disk scattering nuclear light could produce a parallel polarization. In a realistic accretion disk picture, the inner region of the disk itself produces much of the luminosity and partially plays the part of the 'nucleus' in these statements. One important consequence of a geometrically and optically thick torus surrounding the nucleus of a Seyfert 2 galaxy is that our view of the nucleus would be blocked. This may account for

Table 1 Polarization alignments among Seyfert galaxies

Object	Type	Optical polarization		Radio axis (ref.)	Difference	Comments
NGC4051	1	0.59±0.07%	77±3°	78±5°*	1±6	
NGC4151	1	0.57±0.04%	85±2°	83±5° (13)	2±5	
3C390.3	1	1.39±0.21%	153±4°	143±5° (14)	10±6	
NGC3227	1	1.77±0.09%	126±2°	146°† (12)	20†	Map unclear; sometimes misclassified as Seyfert 2, but see ref.15
Mrk 231	1	2.87±0.08%	95±1°	116°† (12)	21†	Unusual spectrum; difficult map
NGC2992	2	3.32±0.18%	33±2°	160±10°*	53±10	
Mrk 3	2	1.78±0.07%	145±1°	86±5° (16)	59±5	
NGC1068	2	2.11±0.02%	98.4±0.3°	18±5° (13)	80±5	
Mrk 348	2	1.80±0.22%	85±3°	167±5°‡	82±6	VLBI map
Mrk 463	2	4.18±0.16%	86±1°	2±5 (12)	84±5	
Mrk 34	2	1.01±0.08%	62±6°	154±5°*	88±8	

* J. S. Ulvestad and A. S. Wilson, in preparation.

† Problematical objects.

‡ S. Neff and A. G. de Bruyn, in preparation.

the inability of the observed UV/optical continuum of NGC1068 to power the observed emission lines and thermal dust radiation⁴.

Thick-disk polarizations would be expected to be higher, and this is seen in both the radio galaxies and this Seyfert sample. The reasons are that thick disks would block our direct view of the unpolarized nucleus and that thin disks might intercept relatively little light because of limited optical depth and limited solid angle from the point of view of the continuum source (see the data of Serkowski¹⁰ for possible galactic analogues). As Seyfert 2 galaxies generally have more dilution of the nuclear light by unpolarized starlight than do Seyfert 1 galaxies, the nuclear polarizations of those shown here are intrinsically substantial and much higher than the Seyfert 1 galaxies. For example, in the case of Mrk 348, Koski¹¹ found that 90% of the H β -region continuum entering a 2.7×4.0 arc s aperture is starlight, so the intrinsic nuclear polarization is quite high. Such a high polarization requires an axisymmetric geometry with considerable scattering optical depth. (Polarization by dust transmission would tend to be perpendicular to the host galaxy, but the Seyfert host galaxy axes are apparently unrelated to the axes of the associated radio sources¹². Besides, such high transmission polarizations would be accompanied by enormous extinction and reddening.)

Finally, if these thin and thick scattering disks are accretion disks, they may result from subcritical and supercritical accretion, respectively.

We conclude that all active object classes having stable optical polarization position angles, including Seyferts, shown polarization alignments. Seyfert 1 polarizations tend to be parallel to the radio axes while Seyfert 2 polarizations tend to be perpendicular. These alignments might be due to thin and thick scattering disks, respectively, and so represent a fundamental difference between the spectroscopic Seyfert galaxy classes.

I thank J. S. Miller for encouragement; J. S. Ulvestad for helping to locate the radio data; J. S. Ulvestad, A. S. Wilson, S. Neff and A. G. de Bruyn for allowing me to make measurements from their unpublished maps; and P. G. Martin and collaborators for permission to quote their measurements before publication. The National Radio Astronomy Observatory is operated by Associated Universities, Inc. under contract with the NSF.

Received 3 December 1982; accepted 9 March 1983.

1. Stockman, H. S., Angel, J. R. P. & Miley, G. K. *Astrophys. J. Lett.* **227**, 55–58 (1979).
2. Antonucci, R. R. J. *Nature* **299**, 605–606 (1982).
3. Thompson, I. B., Landstreet, J. D., Stockman, H. S., Angel, J. R. P. & Beaver, E. A. *Mon. Not. R. astr. Soc.* **192**, 53–74 (1980).
4. Miller, J. S. & Antonucci, R. R. J. *Astrophys. J. Lett.* (submitted).
5. Martin, P. G., Thompson, I. B., Maza, J. & Angel, J. R. P. *Astrophys. J.* **266** (1983).
6. Mathewson, D. S. & Ford, V. L. *Mem. R. astr. Soc.* **74**, 139–182 (1970).
7. Angel, J. R. P. & Stockman, H. S. *A. Rev. Astr. Astrophys.* **18**, 321–362 (1980).
8. Angel, J. R. P., Stockman, H. S., Woolf, N. J., Beaver, E. A. & Martin, P. G. *Astrophys. J. Lett.* **206**, 5–9 (1976).
9. Schmidt, G. D. & Miller, J. S. *Astrophys. J.* **240**, 759–767 (1980).
10. Serkowski, K. *Astrophys. J. Lett.* **179**, 101–106 (1973).
11. Koski, A. T. *Astrophys. J.* **223**, 56–73 (1978).
12. Ulvestad, J. S., Wilson, A. S. & Sramek, R. A. *Astrophys. J.* **247**, 419–442 (1981).
13. Wilson, A. S. & Ulvestad, J. S. *Astrophys. J.* **263**, 576–594 (1982).
14. Harris, A. *Mon. Not. R. astr. Soc.* **158**, 1–11 (1972).
15. Osterbrock, D. E. *Astrophys. J.* **215**, 733–745 (1977).
16. Wilson, A. S., Pooley, G. G., Willis, A. G. & Clements, E. D. *Astrophys. J. Lett.* **237**, 61–64 (1980).

²⁵²Cf plasma desorption in ion implanted mica

M. Maurette*, A. Banifatémi†, S. Della-Negra† & Y. Le Beyec†

* Laboratoire René Bernas, and † Institut de Physique Nucléaire, 91406 Orsay, France

²⁵²Cf plasma-desorption mass spectrometry was first used¹ to determine the masses of heavy molecules of biological interest^{2,3}. In this technique ²⁵²Cf fission fragments penetrate a thin deposit of material (10⁴ cm⁻²) inducing the desorption of neutral and ionized species. The ionized species are accelerated by an external electrical field and their mass analysed using a time-of-flight (TOF) mass spectrometer run in the coincidence mode with each fission fragment. We have tried to assess the analytical and/or simulation potential of heavy-ion stimulated desorption of ions (HISD) in other fields ranging from the so-called 'gas-grain' interactions in astrophysics to the complex ageing of surfaces exposed to strongly ionizing environments in fast neutron fission reactors and fusion reactors. We outline here the potential of HISD for investigating ion implantation effects in insulators, and report a very strong enhancement in the HISD of ionized species from ion implanted mica. This enhancement, which generates heavy-ion clusters up to mass ~500 AMU, grows around a critical fluence of implanted ions, and originates from two distinct types of radiation damage defects.

The TOF spectrometer used for this investigation has been described elsewhere⁴. Both the positive and negative ion mass spectra were obtained with a very low fluence of fission fragments (FF) (~5×10⁴ FF cm⁻²). In order to obtain meaningful and reproducible comparisons between the mass desorption spectra generated by a series of artificial treatments, we replaced the thin deposits previously investigated by the perfect cleavage plane of muscovite mica, which is one of the smoothest and best-defined surfaces available. This perfect cleavage plane also allows large (~100 cm²) but thin (~2-μm thick) sheets of mica to be prepared from the same single crystal with homogeneous characteristics. For the present work we selected a few sheets

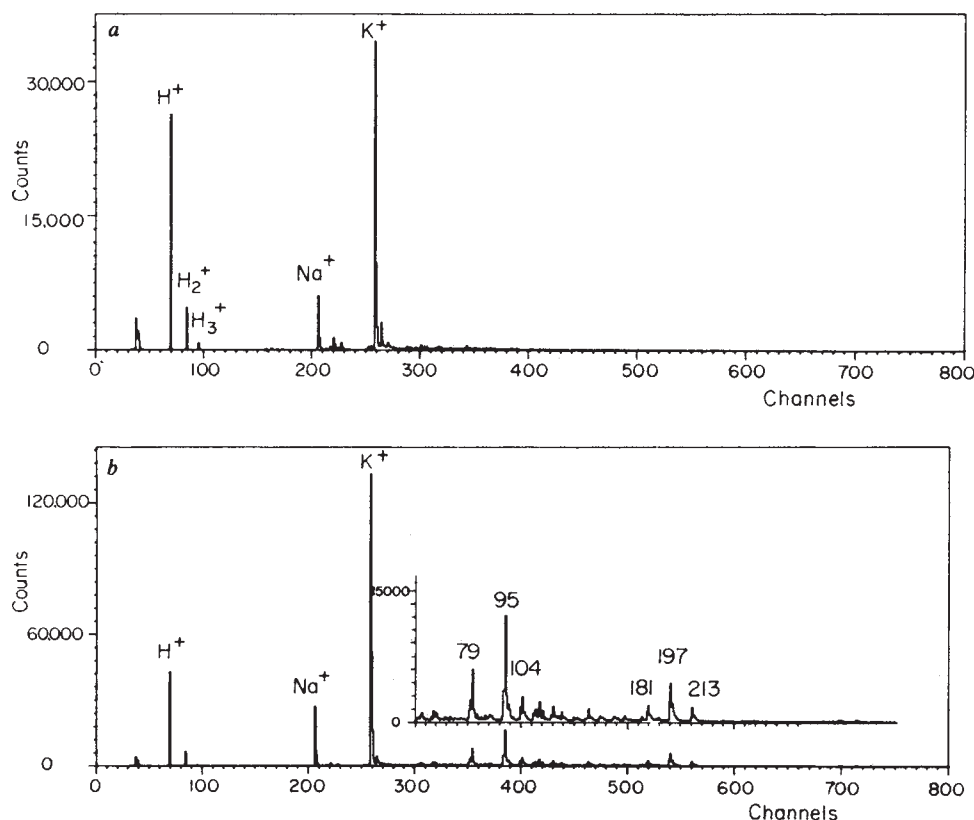


Fig. 1 *a*, Time-of-flight mass spectra of the positive secondary ions desorbed by a low fluence (10^{16} FF cm^{-2}) of fission fragments incident on a 2- μm thick sheet of unirradiated mica. *b*, Similar mass spectra obtained for fission fragments incident on a 2- μm thick sheet of mica implanted with 10^{17} D_2 cm^{-2} . The upper right corner plot represents a 4-fold magnified part of the lower spectra, which can be directly compared with the spectra reported in *a*, and which illustrates the large enhancement of the desorption peak intensities on ion implantation.

with thicknesses ranging from 2 to 10 μm . From each sheet we cut ~ 1 - cm^2 sized targets each with an 'identical' surface and a constant thickness. These sheets of mica were also used to slow down the fission fragments. But mica is a unique nuclear track detector in which the individual tracks of heavy ions can be individually visualized down to a very low energy of ~ 0.2 keV AMU^{-1} (refs 5, 6), when they reduce to a tiny blob of radiation damage with a size of a few 100 $\text{\AA}$. For this reason the characteristics of tracks in mica, including their atomic structure^{7,8}, thermal stability, chemical reactivity and so forth have been investigated in much greater detail than they are in other nuclear track detectors. Note that for the present work, the tracks of the low-energy nuclei used for building up a radiation-damaged layer on mica contain two distinct types of atomic defects; point defects; and extended defects, with the extended defects being much more chemically reactive and thermally stable than the points defects. Consequently thermal annealing runs of the ion-implanted mica should help to demonstrate the role of these two distinct types of defects in plasma desorption.

First we irradiated mica with increasing fluences, $\phi(Z)$, of 1 keV AMU^{-1} ions including D_2^+ , He^+ , $^{13}C^+$, $^{20}Ne^+$ and $^{133}Cs^+$. Our results, which are illustrated in Fig. 1*a, b* are:

(1) The desorption yields $Y(\phi)$ increase sharply around a critical ion fluence, $\phi_D(Z)$, which depends on the atomic number of the implanted ions and then reach a saturation level and/or a maximum at a fluence $\phi_S(Z) \propto 3-5 \phi_D(Z)$. The values of $\phi_D(Z)$ ($\sim 10^{16}$ cm^{-2} for α particles) are similar to the amorphizing fluences $\phi_A(Z)$ required to form an amorphous coating of radiation-damaged material on silicates such as mica, and which has already been observed in transmission electron microscopy on micrometre-sized grains.

(2) However, the detailed shape of the growth curve, $Y_i(\phi)$ depends on the nature of the desorbed species. For example, new peaks appear around $\phi_D(Z)$. This in turn corresponds to very large enhancement factors (≥ 100), which are much higher than those found for major peaks already existing in the unirradiated sample. In particular, these new peaks include the two isotopes of lithium and seven families of peaks centred around the 'magic' mass numbers 63, 79, 95, 103, 114, 181, 197, and 213 in the positive ion spectra and heavy clusters of

negative ions with masses up to ~ 500 AMU , and such as $M(Al_2O_3)_n$, $M'(SiO)_n$ and $M''(SiO_2)_n$.

(3) In contrast with mica, the desorption yields noted for a sheet of Makrofol ($C_6H_{14}O_3$) and a foil of oxidized aluminium coated with a thin layer of Al_2O_3 , generally decreased slightly on ion implantation, and this phenomenon was particularly noticeable at high mass numbers (>100 AMU). This shows that the enhanced desorption of ions from mica is not related to contaminant ions introduced during the ion implantation runs, as all targets were simultaneously irradiated in identical sample holders in the same ion beams.

(4) The sharp increase of $Y(\phi)$ around $\phi_A(Z)$ strongly suggests that the point and extended defects that constitute the amorphous coatings are responsible for the enhanced desorption of secondary ions through a lowering of the binding energy requirements. We thus undertook isochronal thermal annealing runs of 2 h up to 500 $^{\circ}C$ to check the role of the point and extended defects (Fig. 2). Indeed the point defects are much less thermally stable than the extended defects in being annealed at 250 $^{\circ}C$, while the extended defects follow a more complex high-temperature annealing stage, which starts at $T \sim 300$ $^{\circ}C$ and finishes at $T \sim 500$ $^{\circ}C$ (ref. 6). In the reference sample we simply noted a steady decrease in the desorption yields of Na and K which reaches a factor of ~ 3 at 500 $^{\circ}C$ (upper solid in curve Fig. 2). Then a striking departure from this simple trend is observed for a 10^{16} ^{13}C cm^{-2} irradiated sample. Now the Na peak shows the onset of the low-temperature annealing as well as the high-temperature stage expected from the extended defects, while K illustrates the simpler trend of following essentially the high-temperature annealing stage (dotted curves in Fig. 2). On the other hand, the magic mass number 79 and 95 only follow a low-temperature thermal annealing compatible with that of the point defects (solid curve in Fig. 2).

(5) The enhanced desorption observed immediately after ion implantation slightly anneals with time. For example, after about 8 months a decrease of $\sim 50\%$ was observed in the desorption-peak intensities of the magic mass numbers in a 10^{16} ^{13}C cm^{-2} implanted mica, whereas a less pronounced decrease was observed for alkali metals.

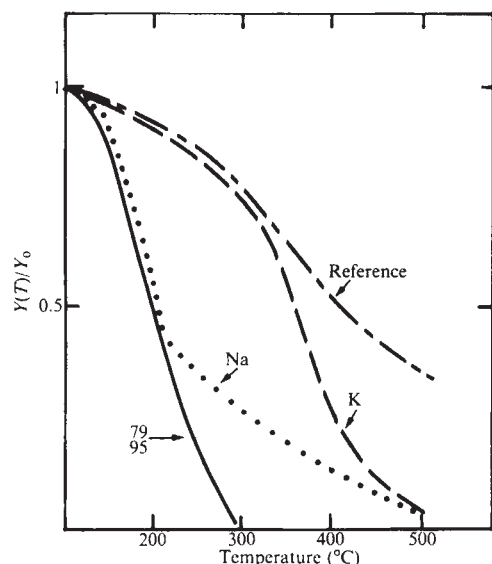


Fig. 2 Variations of the desorption peak intensities with the thermal annealing temperature. The upper dotted line curve refers to the very similar trend noted for both Na and K in the reference sample (unirradiated mica), in which we could not detect the magic mass numbers 79 and 95, yielding the two most intense new desorption peaks on ion implantation. The other curves have been obtained for a $10^{16} \text{ }^{13}\text{C cm}^{-2}$ irradiated sample. We also investigated the annealing curves of the same mass number in a mica implanted with a similar fluence of $1 \text{ keV AMU}^{-1} {}^{20}\text{Ne}$ ions. In this case K shows a much more complex and unexpected behaviour as the desorption yield of this element is now slightly enhanced in the low-temperature annealing range ($T < 300^\circ\text{C}$) and then steeply decreases in the high-temperature range.

(6) For many secondary ions we plotted the decrease of the desorption yields, $Y_i(\Delta)$, with increasing thickness, Δ , of mica. We thus obtained a great diversity of curves, which are mostly confined within an area delimited by two asymptotic trends which scale as $Y \propto S_e$ and $Y \propto S_e^4$, where S_e is the electronic stopping power of the fission fragments. Another interesting result was that the slopes of the $Y(\Delta)$ curves are steeper in the unirradiated sample⁹.

The presence of an external electric field could favour the desorption of alkali metals as well as those of cationized molecules. It would be useful to measure the total sputtering yield of both neutral and ionized species for muscovite mica. These measurements would represent an interesting constraint on the track-formation-plasma-desorption mechanism, and could have interesting implications in other fields. For example, they should contribute to the understanding of the complex ageing of surfaces exposed to strongly ionizing environments to be found in fast neutron fission reactors and in fusion reactors. But they could also be effective as a part of the so-called 'gas-grains' interactions, which attract much attention in astrophysics and planetology, and during which a cosmic dust grain 'aged' while being exposed to various types of ionizing radiations including fluxes of low- and high-energy heavy ions⁹.

Received 13 December 1982; accepted 8 March 1983.

1. Macfarlane, R. D. & Torgerson, D. F. *Science* **191**, 920–925 (1976).
2. Macfarlane, R. D. *Biomedical Mass Spectrometry* **8**, 449–453 (1981).
3. Della-Negra, S., Ginot, Y. M., Le Beyec, Y., Spiro, M. & Vigny, P. *Nucl. Instrum. Meth.* **198**, 159–163 (1982).
4. Le Beyec, Y., Della-Negra, S., Deprun, C., Vigny, P. & Ginot, Y. M. *Rev. Phys. Appl.* **15**, 1631–1637 (1980).
5. Borg, J. thesis, Orsay Univ. (1980).
6. Borg, J. *Nucl. Track. Determin.* (in the press).
7. Dartyge, E., Duraud, J. P., Langevin, Y. & Maurette, M. *Phys. Rev. B* **23**, 5213–5229 (1982).
8. Albrecht, D., Armbruster, P., Roth, M. & Spohr, R. *Nucl. Tracks, Suppl.* **3**, 55–? (1982).
9. Chaumont, J., Della-Negra, S., Lebeyec, Y. & Maurette, M. *Nucl. Instrum. Meth.* (submitted).

Sidescan sonar recognition of subaqueous landforms in Loch Earn, Scotland

John McManus & Robert W. Duck

Department of Geology, The University, Dundee DD1 4HN, UK

Cores obtained during sedimentological studies of British lakes reveal the presence of disturbed layering in slope face and slope foot areas, but undisturbed materials on areas of flat floor. Smith¹ and others have suggested that these features may be due to slumping of the unconsolidated deposits. The abundance of slumping indicated by cores in many lake basins suggested that slope failure and land slipping are very widespread. We now report that sidescan sonar surveys of the slopes of Loch Earn, Scotland, show the presence of over 100 examples of subaqueous slumps which appear to fall into three morphological types. Pinger surveys confirm sediment thickness variations, supporting suggestions of sliding in such areas.

Sedimentological and stratigraphical studies of lakes require not only the collection of surface sediment samples, but also analysis of cores. Whilst cores obtained from the flat central zones of lake basins may yield complete sequential records from continuous accumulation of sediments, many obtained from the marginal or slope foot zones yield interrupted, partly repeated or partially inverted sedimentary successions (see, for example, ref. 1). Only cores showing undisturbed sequences are considered valuable records for palaeoecological or palaeomagnetic purposes, so that those containing evidence of disturbance are classed as 'bad' cores and rejected (W. Pennington, personal communication). Some cores from flat bed areas also contain disturbed sequences.

The palaeoenvironmental histories established from lake sediments are therefore entirely derived from the flat bed areas. However, the proportion of a lake basin occupied by sub-horizontal bed is often small, and in sedimentological terms the processes active within the deposited lake sediments may be dominated by sliding and slope failure. The frequency of occurrence of disturbed sediment sequences in the marginal zones suggests that systematic 'slumping' of material on basin slopes is a normal sedimentological process in lakes.

Bathymetric surveys by Grant Wilson² and Murray and Pullar³ using lead-lining and by Al-Ansari and McManus⁴ using recording echo-sounders agree that Loch Earn ($56^\circ 23' \text{ N}$, $4^\circ 7' - 4^\circ 17' \text{ W}$; maximum depth 87 m) consists of a 10.5 km long, 1 km wide, steep-sided ($8^\circ - 18^\circ$) but largely flat-floored, east-west trending basin which shallows eastwards. Echograms indicate that the basin margins are often lined with terrace-like ledges, and mounds which lack lateral continuity occur at varying depths.

Pinger surveys reveal over 23 m of layered sediments in the deep flat-bed areas, normally resting on tills, thinning to about 2 m on the basin flanks. Locally, 1–2-m high mounds of layered but apparently disturbed sediments occur below slope segments devoid of lake floor muds. Similar features occur at the slope foot, but most lie on the slopes.

Surveys using sideways scanning sonar confirm the presence of numerous surficial structures on the lake floor slopes. In many places sub-continuous shore-parallel features (Figs 1–3) extending to water depths of about 5 m occur in the form of step-like sequences. Up to eight of these are present at any one site in cobble and pebble deposits. These resemble water-level marks and are thought to have formed through wave activity at times of depressed water levels which occur during prolonged dry periods.

Further offshore the sonographs reveal three types of surface features developed in the lacustrine muds.

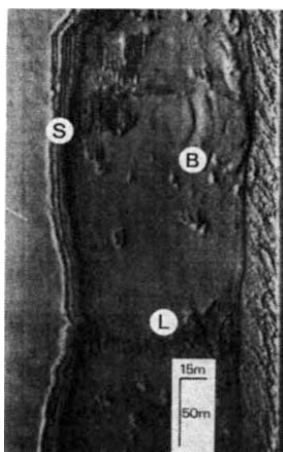


Fig. 1 Sonograph of the bed of Loch Earn. The two straight and parallel lines along the margin of the sonograph indicate the traverse line of the survey vessel. The profile of the lake bed beneath the vessel is shown by the adjacent continuous line. The scan is upslope to the lake shore. Features recognized are shore-parallel steps (S), a linear downslope scour (L) and a zone of blister-like slides (B). The concave, downslope-facing line at B marks the upper limit of the principal slide.

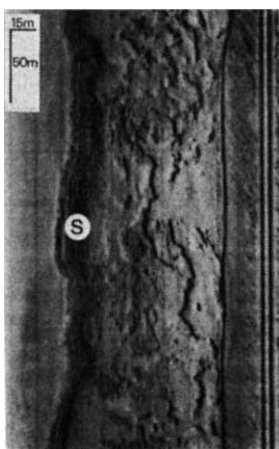


Fig. 2 Sonograph showing shore-parallel steps (S) and a swarm of blister-like slides. Details as for Fig. 1.

(1) Linear downslope scours, localized off small influent streams and drainage culverts (Fig. 1), are erosional channels in which isolated boulders and cobbles are recognized. Narrow at the stream outfall, the scours broaden to 40 m some 200 m offshore. Fan-like sediment accumulations occupy the lower reaches of the shallow gulleys and extend onto the deep basin floors.

(2) Blister-like slides appear as extensive but discontinuous clusters draping the central and lower slopes (Figs 1, 2). Not present at all sites, these locally occur in large numbers. The individual slides vary from 5 m to over 50 m in length but may extend for over 100 m parallel to the shore. It is suggested that these features, which generally have low relief (1–2 m) and which locally exhibit several phases of generation, represent downslope mass movement of deposited muds. The movement appears to have been slow and there is no evidence that anthropogenic disturbances such as mooring pits have served to trigger their formation.

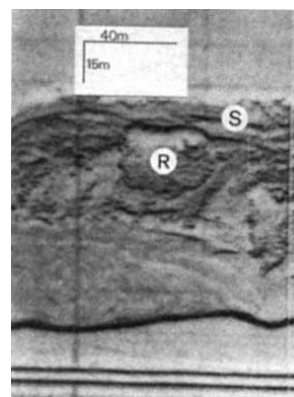


Fig. 3 Sonograph showing shore-parallel steps (S) broken by a rotational slide (R). Details as for Fig. 1.

Rotational slides, apparently analogous to subaerial rotational landslips, have been recognized at a few localities (Fig. 3). They exhibit steep back walls, broken central sectors and apparently raised frontal or lower segments. Unlike the blister slides, these small (30–40 m wide) structures may extend upwards to include failure of the lower suites of shore-parallel steps. In their extreme form of development, collapse of beaches and terrace lines at headlands on the southern shore has produced chaotic spreads of boulders, cobbles and blocks of sediment across the entire basin margin, as though relatively swift failure or catastrophic collapse was responsible.

Most of the features recorded here occur below present day wave base. The 5 m deep shallows at the only outflow from Loch Earn preclude natural drainage to below this level by overland flows. Because the entire catchment area of the loch receives more than 1,800 mm of precipitation per yr (ref. 5) it is most unlikely that lake surface level will ever have fallen substantially below the outflow sill. The blister-like slides all develop at greater depths, suggesting that wave activity is not responsible for their formation. It is recognized that many factors may contribute to the development of the features identified here and the processes involved are currently being investigated. The subaqueous landforms are not confined to Loch Earn and many have been subsequently recognized in sidescan sonar surveys of nearby Loch Lubnaig.

Although the use of sonar devices in marine^{6,7} and estuarine^{8,9} studies is well established, their deployment in lakes for the assessment of landforms has not been recorded previously. From the features observed and reported here it is evident that sidescan sonars have much to contribute in the study of lacustrine sedimentation.

We thank the Marine Geophysics Unit of IGS (Edinburgh) for use of their Klein Sidescan Sonar, Dr N. G. T. Fannin for helpful discussion of the records and NERC for financial support.

Received 21 January; accepted 14 March 1983.

1. Smith, A. J. J. *sedim. Petrol* **29**, 447–453 (1959).
2. Grant Wilson, J. S. *Scott. Geogr. Mag.* **4**, 251–258 (1888).
3. Murray, Sir J. & Pullar, L. *Bathymetrical Survey of the Scottish Freshwater Lochs* (Challenger Office, Edinburgh, 1910).
4. Al-Ansari, N. A. & McManus, J. *Scott. Geogr. Mag.* **96**, 105–113 (1980).
5. McManus, J. in *Dundee and District British Association for the Advancement of Science* (ed. Jones, S. J.) 107–124 (1968).
6. Belderson, R. H., Kenyon, N. H., Stride, A. H. & Stubbs, A. R. *Sonographs of the Sea Floor* (Elsevier, Amsterdam, 1972).
7. McQuillin, R. & Ards, D. A. *Exploring the Geology of Shelf Seas* (Graham & Trotman, London, 1977).
8. D'Olier, B. in *Estuarine Hydrography and Sedimentation* (ed. Dyer, K. R.) 57–86 (Cambridge University Press, 1979).
9. Buller, A. T., Green, C. D. & McManus, J. in *Nearshore Sediment Dynamics and Sedimentation* (eds Hails, J. & Carr, A.) 201–249 (Wiley, London, 1975).

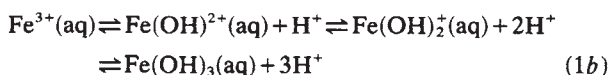
Photo-oxidation of hydrated Fe²⁺—significance for banded iron formations

Paul S. Braterman, A. Graham Cairns-Smith
& Robert W. Sloper

Department of Chemistry, University of Glasgow,
Glasgow G12 8QQ, UK

The Precambrian banded iron formations (BIFs) are the major iron ore sources on the Earth. They consist of extensive iron-rich and iron-poor layers within siliceous sedimentary rocks^{1,2}. The banding has been attributed to variations in the conditions for precipitation of Fe²⁺ in ancient seas. The most favoured precipitating agent is oxygen³⁻⁸; this would lead in the first place to insoluble FeOOH. The variations might then arise from fluctuations in low levels of oxygen in the atmosphere⁹. Similar fluctuations could arise through the *in situ* photosynthetic activity of phytoplankton⁵. Alternatively, oxygen might have been a constant factor, the periodicity arising from a varying supply of iron to the zone of precipitation⁸⁻¹⁰. Another suggestion, that UV photons might have been the precipitating agent, without oxygen¹¹, was based on the conversion of Fe²⁺ to Fe³⁺ by UV light (254 nm) in rather strongly acid conditions^{12,13}. We have now tested this idea by photolysing Fe²⁺ in more realistic near-neutral conditions. We report here that the presence of Fe(OH)⁺ becomes important, greatly enhancing previous estimates of the iron-precipitating power of early Precambrian sunlight, and suggesting that this sunlight would have been a sufficient precipitating agent for the iron found in BIFs.

Reaction (1a) has been well studied in water¹²⁻¹⁴, but only at pH < 3.5, presumably to avoid interference by hydrolysis and precipitation (1b,c).



For us, however, the precipitation was the main object of study. We therefore irradiated open surfaces of more neutral solutions from above in different atmospheres and with variously filtered light from a medium pressure Hg lamp. The reaction was followed by pH drift and analysis for Fe(III). The formation of hydrogen (deuterium from D₂O solutions) was confirmed by mass spectrometry. Solutions were generally 0.56 M NaCl; results in distilled water showed no important differences.

Within the pH range 4.8–6.1 (Table 1) our results were consistent with those found in more acid conditions, Fe(III) initially remaining in solution. On standing, or after more prolonged irradiation, γ-FeOOH was precipitated. Reducing atmospheres (H₂ or CH₄ as opposed to N₂ or Ar) caused no detectable decrease in Fe(III) production. As expected, light of >300 nm was ineffective, being beyond the charge transfer absorption band for [Fe(H₂O)₆]²⁺. The estimated quantum yield (0.06 ± 0.03, see Table 1 legend) was sufficient in principle to account for annual deposition rates in Hamersley BIFs (see calculation in ref. 11).

However, the pH of the oceans 2,000 Myr ago is thought to have been decidedly higher than 6. There is great uncertainty, but estimates of 6.7 (ref. 15), 6.8–7.5 (ref. 5), 7–8 (ref. 8) and ~8 (ref. 16) have been made. We found that on irradiation of Fe²⁺ solutions at pH > 6.5, Fe(III) production increased dramatically, while dark reaction rates remained negligible. We attribute this effect to the formation of [Fe(OH)(aq)]⁺ [equation (2)], for which $K = 3.2 \times 10^{-10} \text{ mol dm}^{-3}$ (ref. 17)

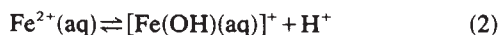


Table 1 Rates of proton release during irradiation of argon-flushed solutions of 10⁻² M Fe²⁺ containing 0.56 M NaCl

pH	$\frac{d[\text{H}^{+}]}{dt} (10^{-9} \text{ mol dm}^{-3} \text{ s}^{-1})$
6.06	6.0 ± 1.0
5.70	6.0 ± 1.0
5.40	6.0 ± 1.0
5.20	5.9 ± 1.0
5.00	5.8 ± 1.0
4.80	5.8 ± 1.0

Irradiation was from above through a suprasil II window on to the open surface of an argon-flushed solution, using a Hanovia UVS 220 medium pressure mercury lamp 25 cm above the solution. Cell depth 30 mm. Projecting into the photolysis cell were a pH combination electrode (Probion, SR217) and a platinum resistance thermometer (Beckman). A 50-mm path length suprasil cell of distilled water was placed between the lamp and photolysis cell. In these conditions the temperature of the photolysis solution remained constant at 22 ± 0.5 °C. Data were recorded using a multichannel system built by W. A. Scott (Scientific Instruments). Argon used in these experiments was scrubbed of residual oxygen by passing over finely divided copper (BASF catalyst R3-11) at 160 °C. Mass spectrometric analysis of the gas showed it to contain less than 1 p.p.m. of oxygen. K₂Fe(SO₄)₂·6H₂O (BDH) in ~10⁻³ M H₂SO₄ was stirred with iron powder (Ventron Puratronic) for 12 h to remove traces of Fe³⁺, after which the material was filtered, recrystallized twice from hot water and dried in a silica gel dessicator. All operations were carried out in a nitrogen glove box. Samples for irradiation were prepared by dissolution of the purified material in argon-flushed 10⁻³ M H₂SO₄. Transfer to the photolysis cell was achieved by using a specially designed vacuum system, after which the pH was adjusted by injecting argon-saturated 0.1 M NaOH. All other reagents used were analytical reagent grade. A series of experiments, using solutions from 8 × 10⁻⁵ to 2 × 10⁻² M Fe²⁺ in cells of pathlength 30, 100 and 220 nm, were performed. Spectra of the photolysis solutions were obtained after withdrawing samples into evacuated spectrophotometer cells. Total Fe(III) was determined colorimetrically using thiocyanate²¹ after acidification of the solutions. Dark experiments showed negligible pH drift and no Fe(III) production. For experiments in which FeOOH precipitated the H⁺, the Fe(III) ratio was found to be 2.2 ± 0.05, while for short irradiation times (<40 min), where there was no precipitation, this ratio was between 1 and 2, consistent with equation (1a, b) and the known¹⁷ hydrolytic behaviour of Fe³⁺(aq). In such experiments there was a slow continued pH drift after irradiation consistent with the dark reaction; equation (1c). The corresponding quantum yield for this reaction, using 2.0 × 10⁻² M Fe²⁺(aq) in 0.56 M NaCl at pH 5.0, was obtained by comparing the initial rate of proton production (in which conditions Fe(OH)₂⁺ predominates¹⁷: the H⁺/Fe(III) ratio should thus be unity) with the Fe³⁺ yield in 0.26 M H₂SO₄ for which Φ = 0.142 (ref. 13), allowances being made in our actinometry for inner filter effects. The quantum yield thus determined was 0.06 ± 0.03. The data in the table show that this yield is reasonably constant within the pH range 4.8–6.06.

Our solutions showed the strong near-UV to visible absorption band described previously¹⁸ for [Fe(OH)(aq)]⁺. This has the general appearance of the Fe²⁺ charge transfer band, but redshifted by ~80 nm and moreover enhanced by more than two orders of magnitude. The spectrum extends beyond 450 nm. Figure 1 shows that it is photoactive. Thus, at an initial pH of >6.5 our iron solutions were sensitive to light at above 400 nm, although as the pH fell towards 6.2, due to acid production [equation (1b)], the reaction at these wavelengths began to turn itself off [reversal of equation (2)].

The 'photosensitization' due to reaction (2) would have greatly enhanced the iron-precipitating power of sunlight. Today's Sun produces ~12 times as much light energy between 300 and 450 nm as it does below 300 nm (ref. 19). (On a photon count this factor rises to ~16.)

The quantum yield of FeOOH from FeOH⁺ was difficult to determine accurately, but lies between 0.01 and 0.05 (Fig. 1 legend) whether in the presence or absence of supersaturated amorphous silica. This quantum yield would have caused a rate of photoprecipitation in Fe²⁺-containing waters having a pH > 6.5 of 150–750 mg cm⁻² yr⁻¹ in the tropics or

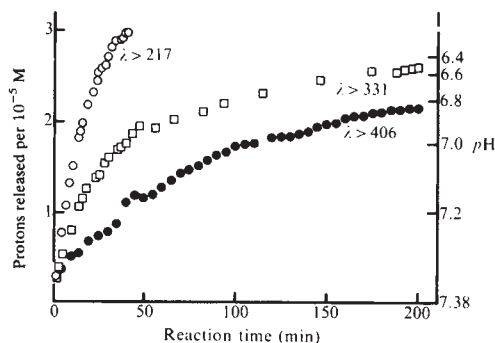


Fig. 1 Wavelength dependence of proton release. Irradiation of 4×10^{-3} M Fe(II) + 0.56 M NaCl was carried out in conditions similar to those described in Table 1. A series of Schott glass filters was used; the cutoff wavelengths given refer to the values for 5% transmission of the filters. The hydrogen ion concentration, $[H^+]$, released was corrected for the considerable buffering effect of $FeOH^+/Fe^{2+}$ using

$$[H^+]_{\text{released}} = [H^+]_t - [H^+]_0 + [FeOH^+]_0 - [FeOH^+]_t$$

$FeOH^+$ concentrations being calculated from $pK_a = 9.5$ (ref. 17). pH changes in a similar dark reaction were small (between -0.02 and $+0.04$ from initial values, over 4 h). In separate experiments using the 331-nm cutoff filter together with a cobalt glass filter, attempts were made to determine the quantum yield for this reaction for 366 nm irradiation. Light intensity was measured using ferrioxalate by the method of Hatchard and Parker²². Colorimetric estimation of Fe(III) after prolonged irradiation confirmed the expected $H^+/Fe(III)$ ratio of ~ 2 . We found quantum yields from Fe(III) or H^+ determinations to be in the range 0.01–0.05, the main source of error here being the high sensitivity of the Fe(II) absorption spectrum to pH.

$50\text{--}250 \text{ mg cm}^{-2} \text{ yr}^{-1}$ at the poles on the basis of today's solar output¹⁹, and allowing for a 50% loss through scattering and absorption by clouds, etc. These values can be compared with estimates for annual iron deposition in Hamersley BIFs of $9\text{--}43 \text{ mg cm}^{-2} \text{ yr}^{-1}$ (ref. 2).

Our calculations depend on the supposition that Fe(III) was efficiently removed by precipitation and sinking from the surface regions of the oceans—otherwise inner filter effects would have reduced yields, perhaps considerably. We do know, however, that the material was precipitated, and the sharp banding, typical of BIFs, suggests prompt deposition; in any case the precipitates formed in our experiments were granular and fell out within hours. Moreover, we have been conservative in not considering absorption of $Fe(OH)^+$ above 450 nm, or any contribution from light below 300 nm which might have been very significant in the absence of an ozone screen and from a much younger Sun with, perhaps, a UV output higher than at present²⁰.

From these considerations, added to the 'overkill' that our calculations imply, it seems that there would have been more than enough photons to have provided the iron-precipitating power for the BIFs. Indeed a major effect of the early Precambrian sunlight on surface waters would have been to keep them generally clear of iron. A top 100 m of ocean that one might imagine initially as being 10^{-4} M in Fe^{2+} would be depleted, according to our calculations, by 1–6 months of sunshine, even in the complete absence of atmospheric oxygen. It follows that renewal of surface iron must have been the factor that controlled both the localities and the periodicities of the precipitation. We would turn, then, to Holland's⁹ arguments for deep ocean waters as having provided the source of iron (and silica) for BIFs, with local periodic upwelling as the efficient cause.

Further details of the photochemical aspects of this work, and a more comprehensive account of the sunlight/upwelling hypothesis for the BIFs, will be presented elsewhere.

We thank Dr T. Baird for electron diffraction studies on precipitates and the NERC for a grant.

Received 14 January; accepted 9 March 1983.

- James, H. L. & Sims, P. K. *Econ. Geol.* **68**, 913–914 (1973); 915–1179 (1973).
- Trendall, A. F. & Blockley, J. G. *Bull. geol. Surv. West. Aust.* **119** (1970).
- Macgregor, A. M. S. *Afr. J. Sci.* **24**, 155–172 (1972).
- Cloud, P. E. *Science* **160**, 729 (1968).
- Cloud, P. E. *Econ. Geol.* **68**, 1135–1143 (1973).
- Schidlowski, M. in *The Early History of the Earth* (ed. Windley, B. F.) 525–535 (Wiley, London, 1976).
- Schopt, J. W. A. *Rev. Earth planet. Sci.* **3**, 213–249 (1975).
- Towe, K. M. *Nature* **274**, 657–661 (1978).
- Holland, H. D. *Econ. Geol.* **68**, 1169–1172 (1973).
- Brock, T. D. *Nature* **288**, 214 (1980).
- Cairns-Smith, A. G. *Nature* **276**, 807–808 (1970).
- Weiss, J. *Nature* **136**, 794 (1935).
- Jortner, J. & Stein, G. J. *phys. Chem.* **66**, 1258–1271 (1962).
- Hayon, E. & Weiss, J. J. *chem. Soc.* 3866 (1960).
- Garrels, R. M. & Perry, E. A. in *The Sea* Vol. 5 (ed. Goldberg, E. G.) 303–336 (Wiley, London, 1974).
- James, H. L. *Econ. Geol.* **49**, 235–203? (1954).
- Baes, C. F. & Mesmer, R. E. *The Hydrolysis of Cations* (Wiley, New York, 1976).
- Ehrenfreund, M. & Leibenguth, J.-L. *Bull. Soc. chim. Fr.* 2494–2505 (1970).
- Weast, R. C. (ed.) *CRC Handbk Chem. Phys.* 61st edn (1980–81).
- Canuto, V. M., Levine, J. S., Augustsson, T. R. & Imhoff, C. L. *Nature* **296**, 816–820 (1982).
- Kolthoff, I. M. & Sandell, E. B. (eds) *Textbook of Quantitative Inorganic Analysis* 3rd edn (Macmillan, New York, 1965).
- Hatchard, C. G. & Parker, C. A. *Proc. R. Soc. A235*, 518–536 (1956).

Postglacial population expansion of forest trees in Norfolk, UK

K. D. Bennett

Sub-department of Quaternary Research, Botany School, Downing Street, Cambridge CB2 3EA, UK

The long life and large size of trees make it unrealistic to use conventional methods of plant biology in the study of their populations¹. However, detailed radiocarbon-dated pollen diagrams offer the opportunity for observing the behaviour of tree populations over long periods of time ($10^2\text{--}10^4$ yr) and to compare phases of expansion with simple mathematical models of population growth². This produces information concerning tree populations which could not otherwise be collected. I report here a pollen-analytical study of the expansion of several tree taxa in the early and mid-postglacial of Norfolk, UK. Exponential and logistic growth models applied to the phases of expansion of these taxa show that their populations double every 35–175 yr.

The investigation was carried out at Hockham Mere, Norfolk (National Grid Reference TL933937), which is an infilled lake basin³, $1,000 \times 750$ m, with a catchment area of $3.8 \times 10^6 \text{ m}^2$. Pollen diagrams from the deposits^{4–6} reveal that there is a continuous sequence of sediments from the Devensian Late-glacial to the present. A core was obtained at the deepest part of the basin for pollen analysis. Pollen counts of $\sim 1,000$ grains were made at 140 levels in the postglacial sequence, at intervals of 1–8 cm (depending on the rapidity of palynological changes)⁷. Pollen concentrations were measured at each level by adding weighed tablets of exotic pollen⁸. The University of Cambridge Radiocarbon Dating Laboratory obtained 23 radiocarbon dates from this sequence. The dates form an internally consistent series and were used to estimate the age [in radiocarbon yr before AD 1950 (BP)] of each pollen sample. The rate of sediment accumulation at the time of deposition of each sample was also calculated. Pollen accumulation rates (PAR) were then obtained from the formula⁹:

$$(\text{PAR}) (\text{grains cm}^{-2} \text{ yr}^{-1}) = \text{pollen concentration (grains cm}^{-3}) \times \text{sediment accumulation rate (cm yr}^{-1})$$

The determination of PAR avoids the constraints imposed by percentage calculations¹⁰, as the values for each pollen type are independent of each other. A pollen diagram from The Mere, Stow Bedon, 3 km north-east of Hockham Mere, shows a quite distinct vegetational history⁷, and reveals that most of the pollen

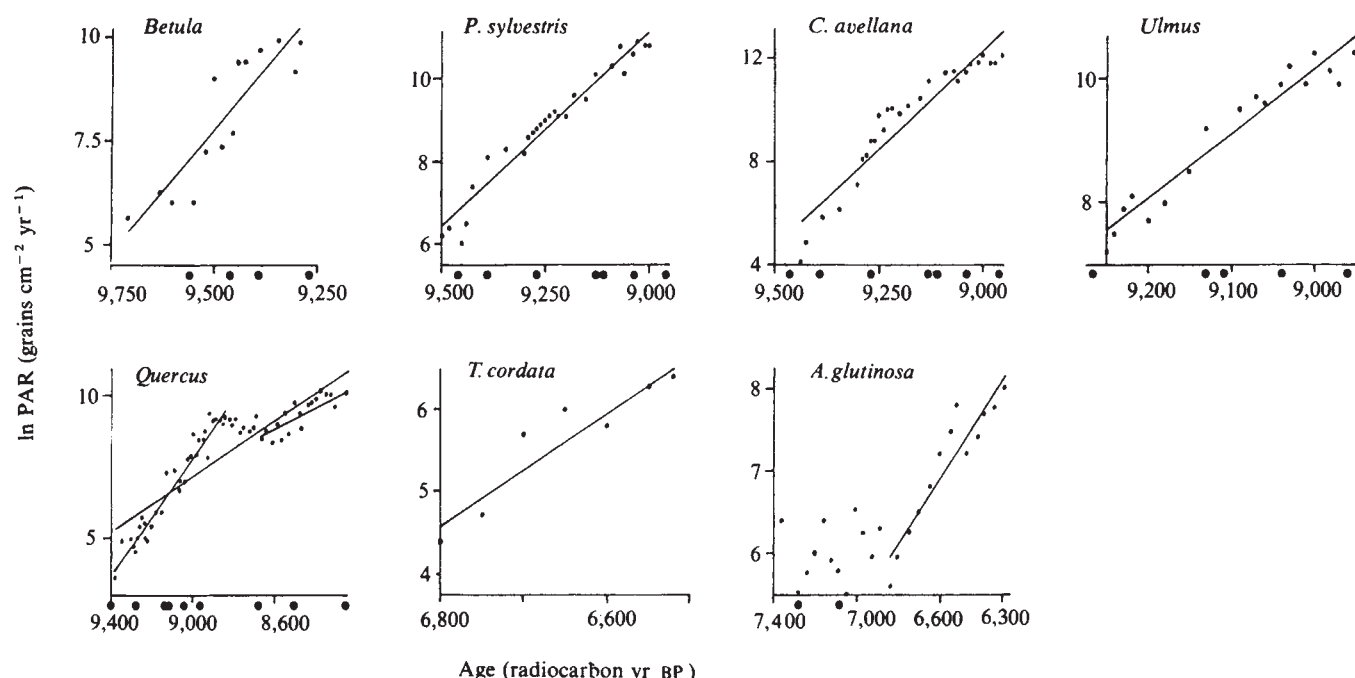


Fig. 1 Plots showing the linear relationship between $\ln \text{PAR}$ ($\ln N$) and sample age for the period of expansion for each of seven taxa. The earliest point on each plot (except that of *A. glutinosa*) represents the beginning of a detectable rise in PAR of the taxon concerned. Points included cover the period until the end of the phase of expansion. Solid dots beneath each plot indicate the position of radiocarbon dates. Other relevant dates are: 10,820 (*Betula*), 7,080 (*T. cordata*) and 6,010 (*T. cordata* and *A. glutinosa*). The curves drawn through the plots are derived from the linear regression analyses. Three curves are included on the *Quercus* plot: one calculated for the complete rise, and one for each of the two phases of increase (see Table 1).

Table 1 Results

Taxon	Correlation coefficient†	r from linear regression $\times 10^{-2}$		Doubling time** (radiocarbon yr BP)		n ††	Duration of expansion‡‡ (radiocarbon yr BP)
		r ¶	Range#	r	Range		
<i>Betula</i> *	-0.89§	1.170	1.550-0.799	59	45-87	14	420
<i>Pinus sylvestris</i> *	-0.98§	0.945	1.030-0.856	73	67-81	27	500
<i>Pinus sylvestris</i> †	0.97§	1.190	1.320-1.060	58	52-66	21	410
<i>Corylus avellana</i> *	-0.95§	1.520	1.730-1.300	46	40-53	27	490
<i>Corylus avellana</i> †	0.97§	1.980	2.190-1.770	35	32-39	23	430
<i>Ulmus</i> *	-0.96§	1.040	1.190-0.886	67	58-78	18	300
<i>Quercus</i> *	-0.87§	0.491	0.564-0.418	141	123-166	59	1,140
<i>Quercus</i> (Phase 1)*	-0.97§	1.050	1.150-0.956	66	60-73	33	550
<i>Quercus</i> (Phase 2)*	-0.80§	0.393	0.547-0.239	176	127-290	18	410
<i>Tilia cordata</i> *	-0.94¶	0.698	1.000-0.394	99	69-176	7	280
<i>Alnus glutinosa</i> *	-0.94§	0.399	0.499-0.300	174	139-231	13	550

* Results of analysis of plots of $\ln N$ against age (equation (2)).

† Results of analysis of plots of $\ln (K - N)/N$ against age (equation (4)).

‡ Correlation coefficients are the opposite sign to that expected from equations (2) and (4) because radiocarbon years are counted backwards from AD 1950. Levels of significance are indicated as follows: § $P < 0.001$; ¶ $P < 0.01$.

¶ Obtained from equations (2) and (4).

95% confidence intervals for r (ref. 13).

** Calculated from r (previous column) using the formula $(\ln 2/r)$.

†† Number of data points included in the analysis.

‡‡ Period of time covered by the points used in the calculations. This is a minimum estimate of the total period of expansion because the time at which expansion began is uncertain.

in the lake is derived from the vegetation of only a small surrounding area. The PAR values are thus representing the former abundances of taxa within the catchment. The proportion of the pollen from outside this area is likely to be small because of the high pollen production of the local forested vegetation.

Since the end of the Devensian cold stage, forest trees have migrated into Britain from their glacial refugia¹¹. The lower sediments at Hockham Mere demonstrate palynologically the sequential arrival of several taxa, and the expansion of their

PAR values to stable levels is seen. Simple mathematical models of population growth were used to discover the duration and nature of these phases of expansion. Exponential increase is the simplest model of population growth, as the change in numbers of individuals over a period of time, t , is proportional to the numbers (N) present¹². This may be written

$$\frac{dN}{dt} = rN \quad (1)$$

where r represents the unrestricted rate of increase per

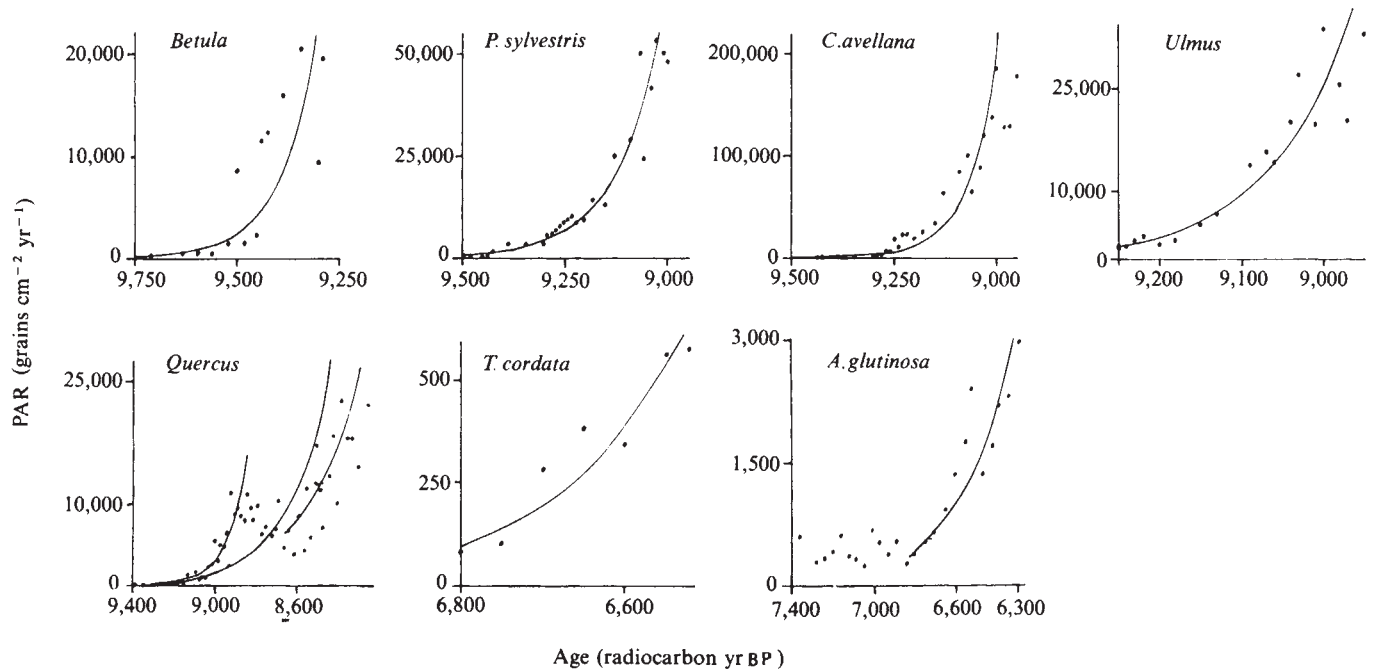


Fig. 2 Plots showing the exponential increase of PAR with decreasing sample age. Plots are given for seven taxa. See Fig. 1 for the position of relevant radiocarbon dates. The curves drawn through the plots are derived from the linear regression analyses. Three curves are included on the *Quercus* plot: one calculated for the complete rise, and one for each of the two phases of increase (see Table 1).

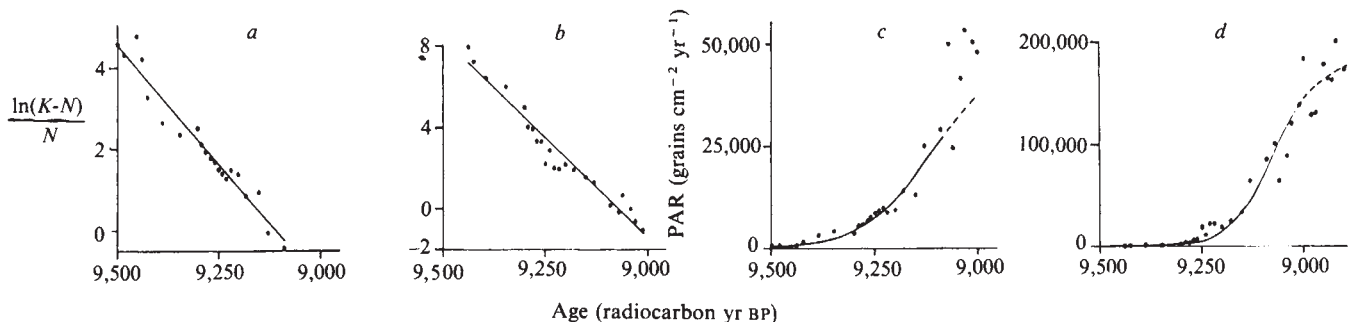


Fig. 3 Plots illustrating the logistic expansion of *P. sylvestris* and *C. avellana*. *a* and *b* show the linear relationship between $\ln(K-N)/N$ and sample age for *P. sylvestris* and *C. avellana* respectively. The value of K for *P. sylvestris* is 47,018 grains $\text{cm}^{-2} \text{yr}^{-1}$ (mean of five levels) and for *C. avellana* 182,344 grains $\text{cm}^{-2} \text{yr}^{-1}$ (mean of 14 levels). *c* and *d* show the logistic growth of populations of *P. sylvestris* and *C. avellana* respectively. See Fig. 1 for the position of relevant radiocarbon dates. The curves drawn through the plots are derived from the linear regression analyses (see Table 1). Values for PAR greater than K cannot be substituted into the logistic growth equation (4). The levels included in the analysis therefore extend only as far as the level before the first point to exceed K . Curves beyond this point are extrapolation and are indicated by dashes.

individual. Equation (1) may be integrated to give

$$N = \exp(rt + a) \quad \text{or} \quad \ln N = rt + a \quad (2)$$

where a is the constant of integration. Equation (2) is the formula of a linear relationship between $\ln N$ and t , with gradient r .

PAR can be considered as a measure of N , so $\ln \text{PAR}$ values for pollen of *Betula*, *Pinus sylvestris*, *Corylus avellana*, *Ulmus*, *Quercus*, *Tilia cordata* and *Alnus glutinosa* were plotted against the radiocarbon time scale for the period when each taxon was expanding (Fig. 1). The *Quercus* curve showed two phases of expansion, so each was treated separately. Correlation coefficients between sample age and $\ln \text{PAR}$ were calculated for each taxon to establish whether the relationship was significant (Table 1). Estimates for r , the gradient, can be obtained from regression analysis. Least-squares linear regressions¹³ of $\ln \text{PAR}$ on age were used because there is clearly an independent variable (age) and a dependent variable ($\ln \text{PAR}$). The estimates for r were then used to obtain population doubling times (Table 1), as these are a convenient way to compare rates of expansion¹. The exponential curves derived from the linear regressions are plotted through the data points (Fig. 2). Thus, the increases in PAR values, and hence the past

tree populations, can be reliably represented by an exponential curve.

The next simplest model of population growth is the logistic growth curve¹².

$$\frac{dN}{dt} = \frac{rN(K-N)}{K} \quad (3)$$

which expresses in the simplest possible way the growth of a population to an upper asymptote K . The integral form of equation (3) may be written:

$$\frac{\ln(K-N)}{N} = a - rt \quad (4)$$

which is the equation of a straight line with gradient $-r$ for which the x and y coordinates are t and $\ln(K-N)/N$ respectively. This model is likely to be more realistic because populations do not continue to expand indefinitely. It was used only, however, for the data of *C. avellana* and *P. sylvestris* because they are simple increases with a larger number of data points than the other taxa. K was obtained by averaging PAR values for each taxon after the completion of its expansion. Correlation coefficients between $\ln(K-N)/N$ and t and linear regressions

of $\ln(K-N)/N$ on age were obtained (Fig. 3, Table 1). The logistic model implies an early rapid rate of increase before a gradual levelling off, whereas the exponential model assumes a constant rate of increase. Thus, the doubling times derived from the logistic model for *C. avellana* and *P. sylvestris* are significantly shorter than those obtained from the exponential model (Table 1). The logistic model represents the increase of PAR values at least as well as the exponential model. Either may be considered as a basis for examining palynological data for information on past tree populations.

The population doubling times vary from 35 to 175 yr, depending on the taxon. These figures must be related to data on the reproductive rates of trees in order to understand what they mean in terms of past population behaviour. The earliest age of flowering of British forest trees is probably in the range 10–50 yr (*P. sylvestris*¹⁴, *Q. petraea*¹⁵, *B. pendula*¹⁶, *Fraxinus excelsior*¹⁷, *U. glabra*¹⁸ and *A. glutinosa*¹⁸), but some specimens of *Quercus* in a closed canopy may not produce fruit until they are 70–80 yr old¹⁸. Since the early postglacial tree populations double in a period of time similar to the age of trees at first flowering, these populations seem to have expanded as fast as they could. The longevity of some tree species may be as low as 100 yr (*Betula*¹⁹, *A. glutinosa*²⁰, *C. avellana*²¹). *P. sylvestris*¹⁴ and *Ulmus*^{19,21} may reach 200–300 yr, while *Quercus*¹⁵ may live much longer. *T. cordata* can survive indefinitely by regrowth from the stem base²². The duration of population expansions (280–550(–1,140) yr; Table 1) is thus greater than the lifespan of most of the taxa concerned. The expansions therefore take place over a period of several generations.

The curve for the expansion of *Quercus* shows two distinct phases, separated by a slight decline (Figs 1, 2). The two phases have been treated separately, and show different rates of increase (Table 1). These may be related to the two palynologically indistinguishable species of *Quercus* in Britain, possibly with the more opportunistic *Q. robur*¹⁹ forming the earlier, more rapid rise, and *Q. petraea* being responsible for the later, slower rise.

Pollen profiles often record a period of low frequencies before a sustained rise is apparent. This is a natural consequence of exponential population increase. However, before the main expansion of *A. glutinosa* there is a long period of low, variable values (Figs 1, 2), which cannot be attributed to early stages of the population build up. In this case, the data suggest a continuous, low population level before the expansion began, which may have been initiated by a climatic change.

Palynological data from Japan²³ and western North America²⁴ on the expansion of *Fagus* and *Pseudotsuga menziesii* suggest that these pioneering and 'sub-climax' arboreal taxa show much faster doubling times than trees of 'climax' forest, such as *Cryptomeria japonica*²⁵. The Hockham Mere data indicate that late immigrants (*Quercus*, *T. cordata* and *A. glutinosa*) expanded at slower rates than those appearing earlier. This may be because interference from established taxa slowed down the rate of increase of new immigrants²⁶, but equally, some taxa may have arrived late because they reproduce at a lower rate and cannot migrate as fast as earlier arrivals. The solution to this problem will lie in greater precision and understanding of the rates and distances of migration as well as the collection of more data on rates of expansion in different environmental situations.

I thank Drs H. J. B. Birks and H. F. Lamb for their comments on earlier drafts of this paper, Dr V. R. Switsur for the radiocarbon dating, Mrs S. M. Peglar for technical help, and Mr J. M. Line for computing assistance. This work was carried out during the tenure of a NERC research studentship. I also thank Dr H. J. B. Birks for help and encouragement.

Received 6 January; accepted 3 March 1983.

1. Harper, J. L. *Population Biology of Plants* (Academic, London, 1977).
2. Watts, W. A. in *Quaternary Plant Ecology* (eds Birks, H. J. B. & West, R. G.) 195–206 (Blackwell, Oxford, 1973).
3. Mosby, J. E. G. *Trans. Norfolk Norwich nat. Soc.* **14**, 61–67 (1935).
4. Godwin, H. & Tallantire, P. A. *J. Ecol.* **39**, 285–307 (1951).

5. Sims, R. E. in *Quaternary Plant Ecology* (eds Birks, H. J. B. & West, R. G.) 223–236 (Blackwell, Oxford, 1973).
6. Sims, R. E. in *The Effects of Man on the Landscape: the Lowland Zone* (eds Limbrey, S. & Evans, J. G.) 57–62 (Council for British Archaeology Research Rep. **21**, 1978).
7. Bennett, K. D. thesis, Univ. of Cambridge (1982).
8. Stockmarr, J. *Pollen Spores* **13**, 615–621 (1971).
9. Birks, H. J. B. & Birks, H. H. *Quaternary Palaeoecology* (Arnold, London, 1980).
10. Faegri, K. & Iversen, J. *Textbook of Pollen Analysis* 3rd edn (Blackwell, Oxford, 1975).
11. Huntley, B. & Birks, H. J. B. *An Atlas of Past and Present Pollen Maps for Europe: 0–13000 years ago* (Cambridge University Press, in the press).
12. Hutchinson, G. E. *An Introduction to Population Ecology* (Yale University Press, 1978).
13. Till, R. *Statistical Methods for Earth Scientists* (Macmillan, London, 1974).
14. Carlisle, A. & Brown, A. H. F. *J. Ecol.* **56**, 269–307 (1968).
15. Jones, E. W. *J. Ecol.* **47**, 169–222 (1959).
16. Longman, K. A. & Wareing, P. F. *Nature* **184**, 2037–2038 (1959).
17. Wardle, P. *J. Ecol.* **49**, 739–751 (1961).
18. Brown, J. & Nisbet, J. *The Forester* (Blackwood, London, 1894).
19. Rackham, O. *Ancient Woodland* (Arnold, London, 1980).
20. McVean, D. N. *J. Ecol.* **41**, 447–466 (1953).
21. Molisch, H. *Die Lebensdauer der Pflanze* (Fischer, Jena, 1929).
22. Pigott, C. D. & Huntley, J. P. *New Phytol.* **81**, 429–441 (1978).
23. Tsukada, M. *Jap. J. Ecol.* **32**, 113–118 (1982).
24. Tsukada, M. *Jap. J. Ecol.* **32**, 159–187 (1982).
25. Tsukada, M. *Jap. J. Ecol.* **31**, 371–383 (1981).
26. Smith, A. G. *Proc. R. Soc. B* **161**, 331–342 (1965).

Uniparental inheritance in a homothallic alga

Karen P. VanWinkle-Swift & Brigitte Aubert

Department of Biology, San Diego State University, San Diego, California 92182, USA

Uniparental inheritance of chloroplast genes is a phenomenon common to both higher plants and eukaryotic algae¹. In the latter case, as exemplified by the heterothallic green alga *Chlamydomonas reinhardtii*, chloroplast gene transmission is controlled by the nuclear mating-type locus^{2,3}. When the stable opposite mating types are crossed, chloroplast markers carried by the mating-type plus (*mt*⁺) gamete are transmitted to the zygote progeny, while the markers (and chloroplast DNA) of the mating-type minus (*mt*[−]) gamete are not, in most cases^{4–7}. Here we report the isolation of a uniparentally inherited erythromycin-resistant mutant (*ery-u1*) of the homothallic (self-mating) alga, *Chlamydomonas monoica*. The inheritance pattern of the *ery-u1* mutant can be interpreted in terms of a general model for homothallism which assumes that each cell of a clonal population carries genetic information analogous to the *mt*⁺ and *mt*[−] alleles of heterothallic species, and that self-mating results from differential gene expression within the clonal population.

Thus far, cytoplasmic (presumably chloroplast) inheritance in *Chlamydomonas* has been studied only in heterothallic species, that is, those consisting of stable opposite mating-types. In the extensively studied species *C. reinhardtii*, and in *Chlamydomonas eugametos*, the presumptive chloroplast antibiotic resistance markers show mating-type-directed uniparental inheritance^{2,3,8,9}. Similar markers in *Chlamydomonas moewusii* apparently may be inherited either biparentally¹⁰ or in a uniparental manner¹¹. At present little is known about the way in which the mating-type locus controls the transmission of chloroplast DNAs during the sexual cycle. Although the mating-type locus of *C. reinhardtii* seems to lie within a region of cross-over suppression¹², an understanding of the molecular nature of the locus, and in particular the differences between *mt*⁺ and *mt*[−] alleles, awaits the successful cloning of this region of the nuclear genome.

Efforts to understand the genetic control of mating-type differentiation in *C. reinhardtii* through studies on mating mutants have been limited by the narrow spectrum of available mutant types and the obvious difficulties encountered when attempting analysis of crosses between mating-defective strains¹³. In contrast, classical genetic approaches using homothallic strains have proved invaluable in unveiling the control mechanisms for mating-type differentiation in the yeast *Saccharomyces cerevisiae*^{14–16}. In an effort to learn more about

Table 1 Uniparental inheritance of the *ery-u1* marker

Cross	Locus	No. of tetrads with segregation ratios (mutant: wild-type) below:				
		4:0	3:1	2:2	1:3	0:4
(a) <i>ery-u1 zym-6</i> × <i>spr-fd-1 zym-1</i>	<i>ery-u1</i>	9*	0	0	0	13
	<i>spr-fd-1</i>	0	0	22	0	0
	<i>zym-6</i>	0	0	14†	0	0
	<i>zym-1</i>	0	0	14†	0	0
	<i>ery-u1 zym-6</i> × <i>spr-fd-1 zym-13</i>	25	0	0	0	16
	<i>spr-fd-1</i>	0	0	41	0	0
	<i>zym-6</i>	0	0	12†	0	0
	<i>zym-13</i>	0	0	12†	0	0
	<i>ery-u1 spr-fd-1 zym-1</i> × <i>zym-6</i>	9	0	0	0	11
	<i>spr-fd-1</i>	0	0	20	0	0
Total for <i>ery-u1</i> locus:		43	0	0	0	40
(b) <i>ery-10 zym-6</i> × <i>spr-fd-1 zym-1</i>	<i>ery-10</i>	0	0	23	0	0
	<i>spr-fd-1</i>	0	0	23	0	0
(c) <i>ery-u1</i> × <i>spr-fd-1</i>	(1) <i>ery-u1</i>	2	0	0	0	0
	<i>spr-fd-1</i>	0	0	0	0	2
	(2) <i>ery-u1</i>	0	0	0	0	17
	<i>spr-fd-1</i>	17	0	0	0	0
	(3) <i>ery-u1</i>	4	0	0	0	4
	<i>spr-fd-1</i>	0	0	8	0	0

a, Crosses showing uniparental inheritance of the *ery-u1* marker in conditions where selection against self-mating by using *zym* markers has been used. b, Inheritance of a mendelian marker (*ery-10*) for erythromycin resistance in a cross using *zym* markers for selection against self-mating. c, Uniparental inheritance of the *ery-u1* marker in a cross without selection against self-mating, but using the *spr-fd-1* (mendelian) marker for identification of inter-strain mating events. (1) Tetrads from zygotes produced by self-mating within the *ery-u1* strain; (2) tetrads from zygotes produced by self-mating within the *spr-fd-1* strain; (3) tetrads from zygotes produced by matings between the *ery-u1* and *spr-fd-1* strains.

* Resistant products were subcloned and 26–52 subclones from each tested for resistance; no evidence of somatic segregation of the *Ery^r* and *Ery^s* alleles was observed, that is, the products seem to be pure for the *Ery^r* marker.

† *zym* markers were scored by complementation testing¹⁸ only in a random sample of tetrads from each cross.

mating-type regulation in *Chlamydomonas*, and to assess the generality of mating-type control over chloroplast inheritance, we have begun studies on the homothallic species, *C. monoica*^{17,18}. We have isolated a mutant strain (*ery-u1*) which is resistant to the antibiotic erythromycin, and which shows uniparental inheritance in crosses where additional complementary recessive lethal markers have been incorporated to select against self-mating events typical of homothallic species (and resulting in transmission patterns superficially identical to uniparental inheritance).

The erythromycin-resistant mutant was isolated after UV irradiation (yielding 37% survival) of a strain carrying a previously induced recessive mendelian mutation (*zym-6*) affecting zygote maturation. Zygotes homozygous for the *zym-6* mutation (produced by self-mating) are nonviable and undergo spontaneous lysis¹⁸. Mutagenized *zym-6* haploid cells were grown in non-selective conditions for 10 days to allow somatic segregation of mutant and wild-type alleles which may be a prerequisite or the efficient recovery of recessive chloroplast mutants¹⁹. After post-mutagenesis growth, 5×10^5 cells (the vegetative progeny of 1.3×10^4 surviving mutagenized cells) were plated on minimal medium supplemented with $100 \mu\text{g ml}^{-1}$ erythromycin. The single mutant recovered (*ery-u1*) was subsequently found to be resistant to at least $1,000 \mu\text{g ml}^{-1}$ of the antibiotic.

Crosses were performed between the *ery-u1 zym-6* strain and two additional maturation-defective strains, each carrying recessive lethal mendelian markers (*zym-1* or *zym-13*) complementary to *zym-6* (ref. 18) and a mendelian marker conferring resistance to the antibiotic spectinomycin (*spr-fd-1*)^{18,20}. The complementing *zym* markers served to select exclusively for inter-strain pairing events as indicated by the regular 2:2 segregation of the *spr-fd-1* marker in all tetrads analysed (Table 1). Complementation testing (see ref. 18 for method) of the tetrad products also verified regular 2:2 segregation ratios for the mutant and wild-type alleles of the *zym-1*, *zym-6* and/or *zym-13* loci. The transmission of the *ery-u1* marker, however, was markedly different. From each cross, the tetrads could be divided into two classes: those in which all four meiotic products were erythromycin-resistant (*Ery^r*), and those in which all four products were erythromycin-sensitive (*Ery^s*) (Table 1a). Thus

the *ery-u1* marker was transmitted in an exclusively uniparental manner.

In *C. reinhardtii*, conditions which alter the normal course of zygote development (maturation) also affect the transmission of chloroplast markers in crosses^{21–23}. Therefore, we have considered the possibility that the *zym* markers—which interfere with the normal maturation process in homozygous mutant zygotes—might show partial expression in heterozygous zygotes and cause aberrations in nuclear meiosis or disrupt normal patterns of cytoplasmic inheritance. The following observations argue against such a hypothesis: (1) the mendelian markers used are unlinked¹⁸ and their regular 2:2 segregation ratios (Table 1) suggest that meiosis is occurring normally; (2) a spontaneous mendelian erythromycin resistance marker (*ery-10*) shows normal 2:2 segregations in crosses involving the same *zym* markers (Table 1b); and (3) when tetrad products carrying only the *ery-u1* or *spr-fd-1* markers are crossed, the *ery-u1* marker continues to show uniparental inheritance in zygotes where 2:2 segregation of the *spr-fd-1* marker identifies matings between the erythromycin-resistant and erythromycin-sensitive strains (Table 1c).

In *C. monoica*, gametogenesis, pair formation and zygote development occur in clonal cultures^{24,25}. Mating pairs are first established via flagellar agglutination (unpublished observations), suggesting that the cells—identical in genetic constitution—have differentiated into two cell types having complementary flagellar differences promoting agglutination. The mating gametes form tandem motile pairs with one member becoming paralysed soon after a cytoplasmic 'thread' is observed to connect the partners. Pair formation and behaviour is thus strikingly similar to that described for heterothallic species, in particular *C. moewusii*, where paralysis is specific to the *mt⁻* gamete²⁶. The similarities in mating interactions observed for *C. monoica* and heterothallic species led us to postulate that *C. monoica* cells (haploid) carry all the necessary information to act as either *mt⁺* or *mt⁻* gametes and that clonal pair formation requires, and is a consequence of, differential gene expression within the population¹⁷. This model can be used to explain the uniparental transmission pattern exhibited by the *ery-u1* mutant by assuming that the nature of the mating-type information

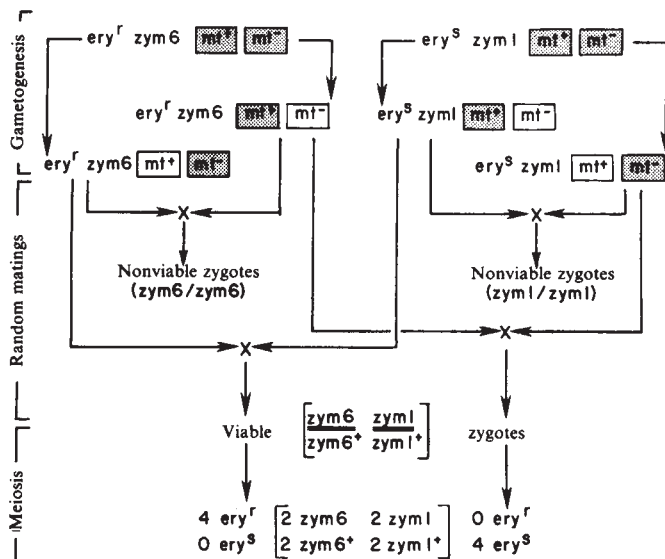


Fig. 1 A model for uniparental inheritance in *C. monoica*. Vegetative haploid cells of *C. monoica* are assumed to carry unexpressed genetic information (shaded boxes) analogous to the mt^+ and mt^- alleles of heterothallic species. In conditions inducing mating (nitrogen and phosphorus limitation in liquid medium), a given strain differentiates into two cell types—one expressing the mt^+ 'locus' (open box) and the other expressing mt^- . Intrastrain pairings thus occur but can be selected against by incorporation of recessive lethal markers expressed only in zygotes (*zym*). If strains carrying complementary selective *zym* markers are mixed and allowed to mate, only heterozygous zygotes produced by inter-strain mating events will be viable¹⁸. When one of these strains carries a cytoplasmic (presumably chloroplast) marker (*Ery*^r), two types of inter-strain matings are hypothesized to occur: those in which the *Ery*^r parent is behaving as the mt^+ gamete, and those in which the *Ery*^r parent is behaving as mt^- . By analogy to *C. reinhardtii*, the chloroplast marker carried by the mt^+ -acting gamete is presumed to be transmitted exclusively to the zygote progeny. In approximately half of the tetrads this will be the *Ery*^r (resistance) allele, while in the remaining tetrads only the *Ery*^s (sensitive) allele will be recovered. In contrast, the mendelian selective markers (*zym-1* and *zym-6*) will segregate 2:2 in all tetrads.

expressed in a given cell also determines whether or not the chloroplast marker carried by that cell will be transmitted after mating (Fig. 1).

Alternatively, uniparental inheritance of the *ery-u1* marker could be explained by a random loss of chloroplast DNA derived from one or the other parent, with the fate of the parental chloroplast DNAs being controlled by genes expressed only after mating. However, it is difficult to envisage a molecular mechanism which could, on the one hand, eliminate the chloroplast DNA of one parent, and on the other hand, ensure the preservation of sufficient chloroplast DNA from the other parent for distribution to the progeny, without invoking differential genetic input from the parent cells.

In *C. reinhardtii*, uniparental inheritance of chloroplast markers is known to be a consequence of the destruction of chloroplast DNA derived from the mt^- gamete^{4,6,8,27}. The control of chloroplast inheritance by the mating-type locus *per se*, or a gene closely linked to it, may be of direct selective advantage. If uniparental inheritance were controlled by a gene unlinked to mating type, random matings within a population could result in the destruction of all chloroplast DNA in many zygotes—a situation likely to be lethal.

Further work with the *ery-u1* mutant should allow critical testing of these models and may further elucidate the mechanism of mating-type differentiation in a homothallic *Chlamydomonas*. The identification of a single mutation which not only interferes with self-mating but also generates highly biased uniparental transmission of the *ery-u1* marker would provide strong support for our working model (Fig. 1). Further-

more, if uniparental inheritance is under mating-type control, then incorporation of the *ery-u1* marker into crosses involving self-sterile strains, such as those described previously¹⁷, may help to distinguish between non-selfing mutants whose self-sterility is due to mating-type-specific defects (crosses should yield only one class of uniparental tetrads with respect to *ery-u1* transmission), and those in which self-sterility is due to non-specific defects (for example, generalized motility mutants) interfering with clonal pair formation (crosses should yield both classes of uniparental tetrads in approximately equal frequencies). The inheritance of the *ery-u1* marker may also prove a useful criterion for identifying environmental conditions or chemical treatments which may lead to an excess production of one expressed mating-type relative to the other in a given population.

This work was supported by a NSF grant (PCM-8010926) to K.V.-S. B.A. was supported by grants from the French government and the Solar Energy Research Institute, Golden, Colorado.

Received 28 January; accepted 11 March 1983.

1. Sears, B. B. *Plasmid* **4**, 233–255 (1980).
2. Sager, R. *Proc. natn. Acad. Sci. U.S.A.* **40**, 356–363 (1954).
3. Chu-der, O.M.Y. & Chiang, K. S. *Proc. natn. Acad. Sci. U.S.A.* **71**, 153–157 (1974).
4. Sager, R. & Lane, D. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2410–2413 (1972).
5. Spreitzer, R. J. & Mets, L. J. *Nature* **285**, 114–115 (1980).
6. Grant, D. M., Gillham, N. W. & Boynton, J. E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6067–6071 (1980).
7. Kuroiwa, T., Kawano, S., Nishibayashi, S. & Sato, C. *Nature* **298**, 481–483 (1982).
8. Gillham, N. W. *Organelle Heredity* (Raven, New York, 1978).
9. McBride, A. C. & McBride, J. C. *J. Phycol.* **11**, 343–344 (1975).
10. Lemieux, C. & Lee, R. W. *Genetics* **94**, s59 (1980).
11. Coleman, A. W. & Maguire, M. J. *Curr. Genet.* (in the press).
12. Gillham, N. W. *Am. Nat.* **103**, 355–388 (1969).
13. Goodenough, U. W., Hwang, C. & Martin, H. *Genetics* **76**, 169–186 (1976).
14. Herskowitz, I. *et al.* in *The Molecular Genetics of Development* (eds Loomis, W. & Leighton, T.) 79–118 (Academic, New York, 1980).
15. Oshima, T. & Takano, I. *Genetics* **97**, 531–549 (1981).
16. Hicks, J., Strathern, J. N. & Klar, A. J. S. *Nature* **282**, 478–483 (1979).
17. VanWinkle-Swift, K. P. & Bauer, J. C. *J. Phycol.* **18**, 312–317 (1982).
18. VanWinkle-Swift, K. P. & Burrascano, C. G. *Genetics* **103**, 429–445 (1983).
19. Lee, R. W. & Haughn, G. W. *Genetics* **96**, 79–94 (1980).
20. VanWinkle-Swift, K. P. *Genetics* **94**, s107 (1980).
21. Gillham, N. W. *Nature* **200**, 294 (1963).
22. VanWinkle-Swift, K. P. *Nature* **275**, 749–751 (1978).
23. Sears, B. B. *Plasmid* **3**, 18–34 (1980).
24. Strehlow, K. Z. *Bot.* **21**, 625–692 (1929).
25. Ettli, H. *Nova Hedwigia* **49**, 1–1122 (1976).
26. Lewin, R. A. *Nature* **166**, 76 (1950).
27. Burton, W. G., Grabow, C. T. & Sager, R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1390–1394 (1979).

Coexistence of neurofilaments and vimentin in a neurone of adult mouse retina

Ursula C. Dräger

Department of Neurobiology, Harvard Medical School, Boston, Massachusetts 02115, USA

Different classes of intermediate filaments are restricted to particular cell types¹. For example, neurofilaments are found only in neurones^{2–6}, whereas filaments that contain the protein vimentin, which were found in some cells of mesenchymal origin and some forms of glia^{4,5,7,8}, are thought to be absent from mature neurones, and present only transiently in early embryonic neurones^{6,9}. However, evidence is presented here of an exception to that rule in the outer plexiform layer of the mouse retina. Double-labelling with antibodies to neurofilaments and vimentin showed that both types of intermediate filaments coexisted in the axonless horizontal cell of that retinal layer, recalling the previous notion that these cells are glial or intermediate between neuronal and glial (reviewed in ref. 10).

Horizontal cells in the outer plexiform layer of the mammalian retina have long been studied by neurofibrillar staining methods (see ref. 10 for a review), which are known to identify neurofilaments^{11–13}. More recently, it has been learned that

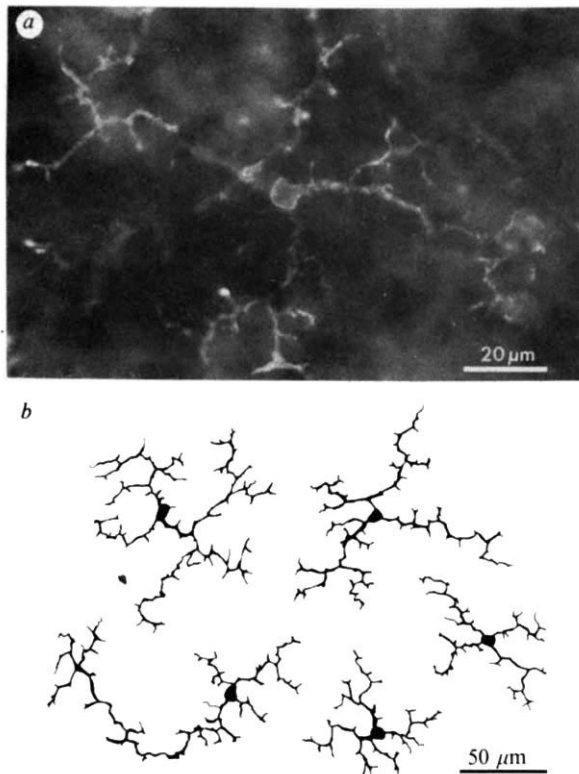


Fig. 1 *a*, Photomicrograph of a microglial cell in a horizontal section through the outer plexiform layer, labelled with an antibody to the macrophage surface antigen Mac-1 (ref. 23). All processes of this cell were contained within the 10-μm-thick section. *b*, Camera lucida drawings of five microglial cells from the outer plexiform layer. As the extended illumination that was necessary for drawing the cells caused some bleaching of the fluorescent label within the field of view, the cells shown here were taken from different fields but were drawn roughly at the spacing seen in the outer plexiform layer.

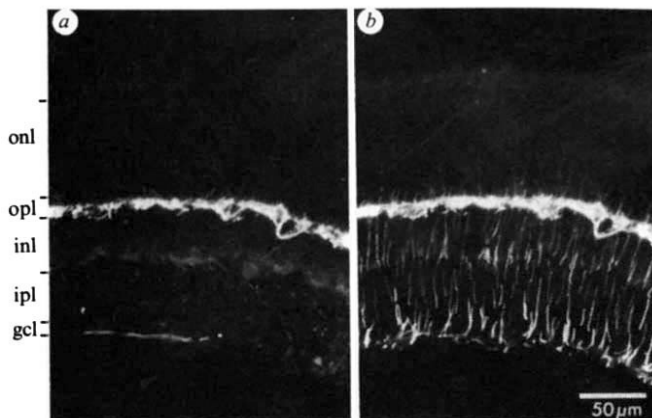


Fig. 2 Cross-section of retina double-labelled with neurofilament antibody RT 97 (ref. 21) in *a* and vimentin antiserum²² in *b*. The binding of the neurofilament antibody was visualized with FITC-labelled goat anti-mouse immunoglobulin fraction (Cappel) and the vimentin antiserum with FITC-labelled goat anti-rabbit antibody (affinity purified; Boehringer-Mannheim). Controls showed both secondary antibodies to be specific for the species against which they were raised. Retinal layers are indicated as: onl, outer nuclear layer; opl, outer plexiform layer; inl, inner nuclear layer; ipl, inner plexiform layer; gcl, ganglion-cell layer. Note the heavy labelling with both antibodies of apparently identical processes in the outer plexiform layer. In addition, the neurofilament antibody stained optic axons in the ganglion-cell layer, some processes in the inner plexiform layer and some weakly neurofilament-positive profiles can be seen at the inner margin of the inner nuclear layer. The vimentin antiserum stained, in addition to the processes in the outer plexiform layer, the radial Müller glia and astroglia in the optic fibre layer.

only one of the two types of horizontal cells found in most mammals, the axonless or A-type horizontal cell, stains by neurofibrillar methods. At an ultrastructural level this cell is unusually rich in neurofilaments^{10,14,15}.

Antisera to the filamentous protein vimentin are known to label the radial Müller glia in the retina^{8,16}, but the labelling in the outer plexiform layer described here has been little studied. Within the mature brain vimentin is thought to be restricted to some types of glia, and in many cases it is associated with the astroglia-specific marker, glial fibrillary acidic protein (GFAP)^{4,5,8,16,17}. No astroglia has been described in the outer plexiform layer, and antisera to GFAP do not label anything in that location (see refs 16, 18 for the rat, and my observations in the mouse). However, another type of glia, microglial cells, with horizontal processes in the outer plexiform layer, have been found in several species^{19,20}. Although microglial cells are not known to be rich in intermediate filaments, it could not be definitely excluded that they might be the profiles in the outer plexiform layer that were labelled with the vimentin antiserum.

Adult C57BL/6J mice were perfused through the heart with 4% paraformaldehyde. The retinas were processed as whole mounts, or were sectioned transversely or horizontally at 10 μm on a cryostat. As neurofilament markers, either RT 97 described by Anderton *et al.*²¹ or R3 (U.C.D., Edwards and Barnstaple, unpublished) were applied; these two monoclonal antibodies recognize strongly the heaviest and weakly the intermediate molecular weight subunits of the neurofilament triplet protein. Both antibodies recognized the same structures in sections of retina, as could be seen in double-labelled specimens with subclass-specific secondary antibodies (Nordic), since RT97 is of the IgG1 and R3 of the IgM class. For vimentin the rabbit antiserum described by Hynes and Destree²² was used. As possible markers for microglia three rat monoclonal antibodies against the macrophage surface antigens Mac-1 (M1/70), Mac-2 (M3/38) and Mac-3 (M3/84) were tested²³⁻²⁵. Antibody binding in the tissue was visualized by indirect immunofluorescence, using fluorescein isothiocyanate (FITC)- and rhodamine isothiocyanate (RITC)-tagged secondary antibodies (Boehringer-Mannheim; Cappel).

In tissue culture microglia/macrophages can be marked with antibodies to immunoglobulin by the presence of Fc receptors on these cells²⁶. In poorly fixed retinas I have seen labelling with secondary antibody to immunoglobulin alone, but the degree of formalin fixation that seemed desirable for reasonable histology largely prevented Fc receptor-specific binding. As an alternative, three macrophage surface markers were tested, but only one of them, an antibody to the antigen Mac-1 (ref. 23), gave staining in these conditions. In transverse sections spidery cells with crenated outlines were seen sparsely distributed throughout the retina, but most frequently in the plexiform layers. In their soma or proximal processes these cells contained chunks of broadly fluorescent material, presumably lipofuscin or remnants of endocytic activity²⁷. The cells in the outer plexiform layer were remarkable in that their processes remained confined to this narrow layer. Their cell bodies, located within the plexiform layer, gave off processes arborizing over an area of 100–150 μm, as was best seen in horizontal sections: Fig. 1*a* shows a photomicrograph and Fig. 1*b* camera lucida drawings of such cells. The five cells in Fig. 1*b* were not actually neighbours (for technical reasons, see Fig. 1 legend), but they were drawn at about the normal spacing seen in the outer plexiform layer. Each cell preserved its own territory at a moderate distance from others; the processes of neighbouring cells never overlapped, but neither were any large spaces left. In all characteristics these cells fitted the description of microglial cells in the retinas of other species, as demonstrated with the del Rio-Hortega, Golgi and ultrastructural methods^{19,20}.

Antibodies to intermediate filaments labelled many more processes in the outer plexiform layer than the macrophage marker. Neurofilament labelling of a cross-section, viewed with the filter set for RITC, is shown in Fig. 2*a*. A band of dense staining can be seen in the outer plexiform layer. In the

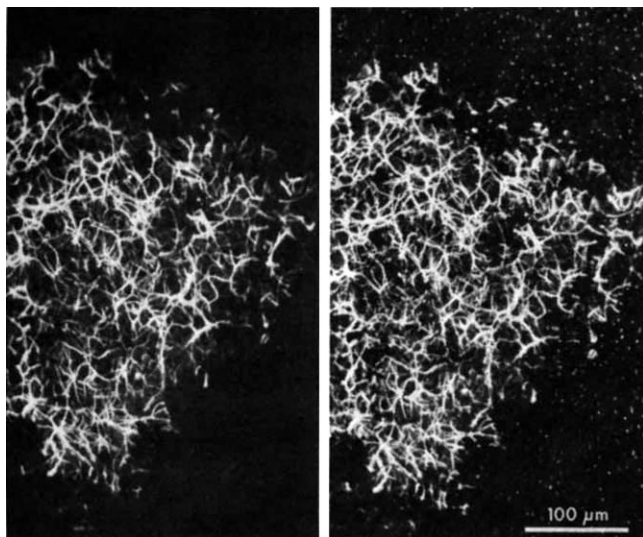


Fig. 3 Horizontal section double-labelled for neurofilaments (left) and vimentin (right) as described in Fig. 2 legend. In the centre of the field the plane of section went through the outer plexiform layer, but passed through the inner nuclear layer elsewhere. In the outer plexiform layer both antibodies labelled identical profiles. In the inner nuclear layer only cross-sections through the vimentin-positive Müller fibres are visible.

ganglion-cell layer a few optic axons are labelled; in addition, although not visible here, neurofilament antibodies may stain large ganglion-cell bodies and some amacrine cells (U.C.D., Edwards and Barnstable, unpublished). The same section, double-labelled with vimentin antiserum and viewed with the filter set for FITC, is shown in Fig. 2b. The radial Müller glia now stands out clearly. The staining of the astroglia in the optic fibre layer, barely distinguishable here, could be better appreciated in horizontal sections (not shown). In the outer plexiform layer the vimentin-positive processes were identical to the processes recognized by the neurofilament antibodies, as was obvious in cross-sections but even more so in oblique sections (Fig. 3). The plane of this section went through the inner nuclear layer, grazing the outer plexiform layer in the central portion of the field shown. The neurofilament antibody (left) labelled only processes in the outer plexiform layer. The vimentin antiserum (right) labelled the same processes, but also the regularly spaced Müller fibres that stand out in the inner nuclear layer. Examination of the labelled processes at high power showed both antigens distributed in identical fashion. Individual filaments could not be distinguished, however, because both antibodies stained the processes solidly. No cells resembling the microglia were labelled by either of the antibodies.

Figure 4 shows camera lucida drawings of single elements from whole mount preparations. These cells had no axon and the branching pattern of the thick dendrites usually lacked an obvious symmetry centre to indicate the location of the cell body that was not labelled. Although drawn apart for clarity in Fig. 4, they were densely packed in the outer plexiform layer, each retinal point covered by several cells. Their shape is reminiscent of the axonless or A-type horizontal cell in the cat^{10,28,29}, and indistinguishable from the same cell type in the rabbit retina, where it has recently been studied by intracellular recordings and injections of tracers^{30,31}. Although similar in lateral dimensions to the microglia, the axonless horizontal cells were dramatically different in shape, branching pattern and spacing.

Whereas most mammals are thought to have an axonless and an axon-bearing type of horizontal cell, both types described in primates have an axon^{28,29,32,33}. Because of this characteristic distribution it was of interest to apply the antibodies to other species. Double-labellings of the rat retina, not surprisingly, showed a similar picture to that in the mouse. In a rhesus macaque, however, processes in the inner retina—optic axons,

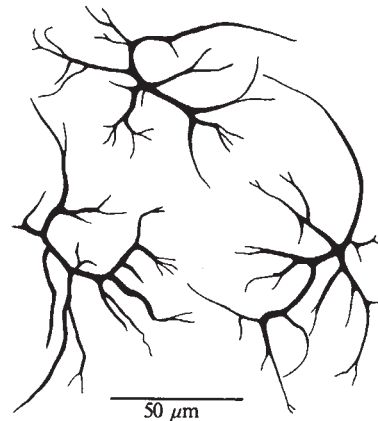


Fig. 4 Three axonless horizontal cells, labelled with RT 97 and traced by camera lucida from retinal whole mounts. In adult mice the cell bodies of these cells were never labelled by neurofilament antibodies or vimentin antiserum. In the developing retina, however, during the period when the dendrites of the horizontal cells slowly assume their horizontal orientation, the cell somata, located in the outermost row of cells in the outer nuclear layer, were labelled with the neurofilament antibodies.

amacrine dendrites—were well labelled with the neurofilament antibody RT 97, but the only labelling in the outer retina was for cell somata outlined at the locations of the horizontal cell bodies in the outermost row of cells of the outer nuclear layer. The vimentin antiserum failed to label the radial Müller fibres in the monkey.

Axonless horizontal cells in mammals receive input from cone photoreceptors³⁴. In certain lower vertebrates, some horizontal cells have been shown to feed back onto cones³⁵, and horizontal cells feed forward into bipolar cells, presumably generating the antagonistic surround response in these cells³⁶. Before such a neuronal function was established through electrophysiological recordings³⁵⁻³⁷, morphological arguments had led to the proposal that horizontal cells might be glial in nature (ref. 38 and reviewed for axonless horizontal cells in ref. 10). The shape of the axonless horizontal cell, with thick dendrites and absence of an axon, extensive electrical coupling between neighbours^{15,30,39} and the lack of unequivocal ultrastructural specializations for the assumed feedback onto photoreceptors, are unusual features of this neurone. The present demonstration of vimentin in this cell, a protein otherwise absent from adult neurones, reinforces the impression that the axonless horizontal cell may share some of the properties of glial cells.

I thank Drs S. Burden, D. H. Hubel and E. Raviola for helpful discussions; Dr R. O. Hynes for the vimentin antiserum; Dr T. A. Springer for the macrophage markers; and Dr J. N. Wood for the neurofilament antibody RT 97. This work was supported by NIH grant EY 01938.

Received 31 December 1982; accepted 14 March 1983.

1. Lazarides, E. *Nature* **283**, 249–256 (1980).
2. Liem, R. K. H., Yen, S.-H., Salomon, G. D. & Shelanski, M. L. *J. Cell Biol.* **79**, 637–645 (1978).
3. Schnitzer, M., Smith, C. & Schoonmaker, G. *Dev. Neurosci.* **1**, 1–14 (1978).
4. Yen, S.-H. & Fields, K. L. *J. Cell Biol.* **88**, 115–126 (1981).
5. Shaw, G., Osborn, M. & Weber, K. *Eur. J. Cell Biol.* **26**, 68–82 (1981).
6. Jacobs, M., Choo, Q. L. & Thomas, C. *J. Neurochem.* **38**, 969–977 (1982).
7. Franke, W. W., Schmid, E., Osborn, M. & Weber, K. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5034–5038 (1978).
8. Schnitzer, J., Franke, W. W. & Schachner, M. *J. Cell Biol.* **90**, 435–447 (1981).
9. Bignami, A., Raju, T. & Dahl, D. *Dev. Biol.* **91**, 286–295 (1982).
10. Wässle, H., Peichl, L. & Boycott, B. B. *Proc. R. Soc. B203*, 269–292 (1978).
11. Gray, E. G. & Guillery, R. W. *J. Physiol., Lond.* **157**, 581–588 (1961).
12. Potter, H. D. *J. comp. Neurol.* **143**, 385–410 (1971).
13. Gambetti, P., Autilio-Gambetti, L. & Papasozomenos, S. C. *Science* **213**, 1521–1522 (1981).
14. Fisher, S. K. & Boycott, B. B. *Proc. R. Soc. B186*, 317–331 (1974).
15. Kolb, H. *J. Neurocytol.* **6**, 131–153 (1977).
16. Bignami, A. & Dahl, D. *Expl Eye Res.* **28**, 63–69 (1979).
17. Dahl, D., Bignami, A., Weber, K. & Osborn, M. *Expl Neurol.* **73**, 496–506 (1981).

18. Dixon, A. S. & Eng, L. F. *J. comp. Neurol.* **195**, 305–321 (1981).
19. Vrabec, F. *J. Neuropath. exp. Neurol.* **29**, 217–224 (1970).
20. Boycott, B. B. & Hopkins, J. M. *Neuroscience* **6**, 679–688 (1981).
21. Anderton, B. H. *et al. Nature* **298**, 84–86 (1982).
22. Hynes, R. O. & Destree, A. T. *Cell* **13**, 151–163 (1978).
23. Springer, T., Galfre, G., Secher, D. S. & Milstein, C. *Eur. J. Immun.* **9**, 301–306 (1979).
24. Ho, M.-K. & Springer, T. A. *J. Immun.* **128**, 1221–1228 (1982).
25. Ho, M.-K. & Springer, T. A. *J. Biol. Chem.* **258**, 636–642 (1983).
26. Raff, M. C. *et al. Brain Res.* **174**, 283–308 (1979).
27. Peters, A., Palay, S. & Webster, H. de F. *The Fine Structure of the Nervous System* (Saunders, Philadelphia, 1976).
28. Gallego, A. in *The Structure of the Eye* (ed. Hollyfield, J. G.) 151–164 (Elsevier, Amsterdam, 1982).
29. Boycott, B. B., Peichl, L. & Wässle, H. *Proc. R. Soc. B203*, 229–245 (1978).
30. Bloomfield, S. A. & Miller, R. F. *J. comp. Neurol.* **208**, 288–303 (1982).
31. Dacheux, R. F. & Raviola, E. *J. Neurosci.* **2**, 1486–1493 (1982).
32. Boycott, B. B. & Kolb, H. *J. comp. Neurol.* **148**, 115–140 (1973).
33. Kolb, H., Mariani, A. & Gallego, A. *J. comp. Neurol.* **189**, 31–44 (1980).
34. Kolb, H. *J. comp. Neurol.* **155**, 1–14 (1974).
35. Baylor, D. A., Fuortes, M. G. F. & O'Bryan, P. M. *J. Physiol., Lond.* **214**, 265–294 (1971).
36. Werblin, F. S. & Dowling, J. E. *J. Neurophysiol.* **32**, 339–355 (1969).
37. Kaneko, A. *J. Physiol., Lond.* **207**, 623–633 (1970).
38. Svætichin, G., Negishi, K., Fatehchand, R., Drujan, B. D. & Selvin de Testa, A. *Prog. Brain Res.* **15**, 243–266 (1965).
39. Raviola, E. & Gilula, N. B. *J. Cell Biol.* **65**, 192–222 (1975).

Single sodium channels from rat brain incorporated into planar lipid bilayer membranes

Bruce K. Krueger*, Jennings F. Worley III* & Robert J. French†

Departments of Physiology* and of Biophysics†, University of Maryland School of Medicine, Baltimore, Maryland 21201, USA

A voltage- and time-dependent conductance for sodium ions is responsible for the generation of impulses in most nerve and muscle cells¹. Changes in the sodium conductance are produced by the opening and closing of many discrete transmembrane channels². We present here the first report of electrical recordings from voltage-dependent sodium channels incorporated into planar lipid bilayers. In bilayers with many channels, batrachotoxin³ (BTX) induced a steady-state sodium current that was blocked by saxitoxin⁴ (STX) at nanomolar concentrations. All channels appeared in the bilayer with their STX blocking sites facing the side of vesicle addition, allowing us to define that as the extracellular side. Current fluctuations due to the opening and closing of single BTX-activated sodium channels were voltage-dependent (unit conductance, 30 pS in 0.5 M NaCl): the channels closed at large hyperpolarizing potentials. Slower fluctuations of the same amplitude, due to the blocking and unblocking of individual channels, were seen after addition of STX. Block of the sodium channels by STX was voltage-dependent, with hyperpolarizing potentials favouring block. The voltage-dependent gating, ionic selectivity and neurotoxin sensitivity suggest that these are the channels that normally underlie the sodium conductance change during the nerve impulse.

Planar bilayers were formed by the method of Mueller *et al.*⁵ across a 0.25–0.5 mm diameter hole in a polystyrene partition separating two chambers, both normally containing 0.5 M NaCl, 10 mM HEPES, 0.15 mM CaCl₂, 0.1 mM MgCl₂, 0.05 mM EGTA, pH 7. The membrane-forming solution contained 33 mg ml⁻¹ phosphatidylethanolamine (bovine brain; Avanti Polar Lipids) and 13 mg ml⁻¹ phosphatidylserine (bovine brain; Avanti) in decane. The current across the bilayer was measured and command voltages were applied to the bathing solutions via a pair of Ag/AgCl electrodes. The side to which vesicles were added was designated the *cis* side; the opposite (*trans*) side was held at virtual ground. The area-specific resistance of the bilayer before addition of biological material was ~10⁸ Ω cm²; the cutoff frequency of the current-to-voltage converter was ~80 Hz. All experiments were performed at ambient temperature (22–25 °C).

Membrane vesicles were prepared from rat brain homogenates as described by Krueger *et al.*⁶. Briefly, fresh rat forebrains were homogenized in isotonic sucrose, using a cavitating

tissue disrupter. Low-speed (1,000g) and intermediate-speed (10,000g) pellets were discarded and the high-speed (100,000g) pellet was suspended in 0.4 M sucrose and stored at -77 °C for up to 6 months before use. The specific activity of ³H-STX binding was enriched about fivefold in this fraction compared with the crude homogenate; at least 60% of the binding sites were recovered in this fraction. We have successfully incorporated sodium channels from at least 12 different membrane preparations obtained by this method.

Incorporation of sodium channels was achieved by the 'fusion' method^{7–9}. Addition of membrane vesicles to the *cis* chamber, with 0.6 μM BTX present on the *trans* side, resulted in stepwise increases in conductance. After 5–30 min, the conductance of the membrane increased by 30–1,000 pS. Before addition of membrane vesicles, the conductance was 10–20 pS. Both BTX and vesicles were required to produce the conductance increase. Figure 1a shows the current across a planar bilayer following addition of membrane vesicles (10 μg ml⁻¹) as described above. About 15 pA flowed across the membrane at ±30 mV: a steady-state, voltage-independent sodium conductance is not surprising because, in the presence of BTX, sodium channels in nerve close only when the membrane is hyperpolarized beyond -90 mV (refs 10, 11 and see below). Addition of 0.3 μM STX to the *cis* side resulted in almost a complete block of the membrane conductance. In this experiment, the conductance in the presence of STX was nearly identical to that seen before incorporation of membrane vesicles. In general, the number of channels incorporated and the rate of incorporation are dependent on the concentration of membrane protein added to the *cis* side. At high concentrations (10 μg ml⁻¹) several channels usually appear in the bilayer within 5 min. In order to obtain only a single channel or very few channels, very low protein concentrations (0.2 μg ml⁻¹) are used and longer periods (up to 30 min) are required. Channels can usually be studied for at least 30 min after incorporation; in some cases a single membrane with one or a few channels has been studied for 2 h, with repeated changes in ionic conditions by perfusion of the *cis* chamber.

Figure 1b shows the effect of varying concentrations of STX (*cis*) on the conductance of the bilayer. Half-maximal block of the conductance occurred at about 5 nM STX (at +30 mV), which is similar to the *K_i* for block of macroscopic sodium currents in nerve by STX^{12–14} and to the *K_d* for binding of radiolabelled STX^{6,15}. Block of the conductance by STX was completely reversed by perfusing the *cis* chamber with toxin-free solution. Addition of up to 1 μM STX to the *trans* side did not affect the membrane conductance. This sidedness of STX block indicates that the channels were incorporated into the bilayer with their extracellular sides facing *cis*.

Figure 2a shows a current record taken after exposure of the bilayer to membrane vesicles in the presence of BTX, and subsequent block by 0.3, 1.3 and 5 μM STX (*cis*). The stepwise fluctuations in the presence of STX seemed to reflect the blocking and unblocking of individual sodium channels, because the mean duration of the conducting state was substantially shortened by raising the STX concentration on the *cis* side. Measurement of the mean times in the zero- and one-channel conductance levels at +30 mV, indicated that the off-rate and on-rate constants for STX block were 0.05 s⁻¹ and 1.1 × 10⁷ s⁻¹ M⁻¹, respectively, implying a dissociation constant of 4.5 nM (compare with Fig. 1b). In many experiments, with very few channels incorporated, discrete, stepwise conductance fluctuations of similar magnitude were also observed in the absence of STX. Examples of these single-channel fluctuations can be seen in Figs 2a, b and 3 (insets).

The unexpected observation that block of BTX-activated sodium channels by STX was voltage dependent is shown in Fig. 2b. After addition of 0.32 μM STX to the *cis* side of a membrane containing seven sodium channels, the channels remained conducting about 5% of the time at -30 mV but only 1% of the time at +30 mV. There was an even more dramatic difference between the percentage block at -60 mV and that

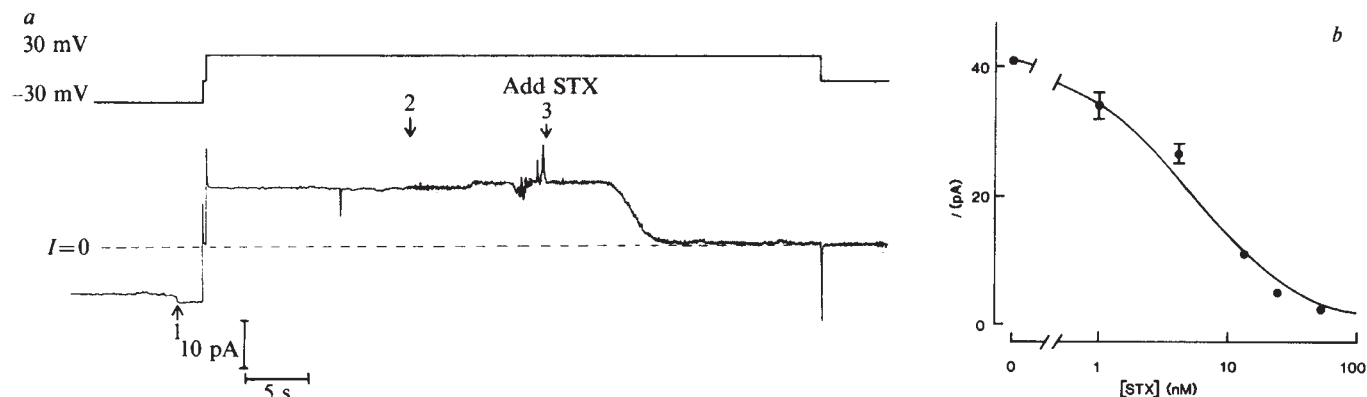


Fig. 1 *a*, Block of conductance by STX. Rat brain membrane vesicles were incorporated into the planar bilayer as described in the text, with 0.5 M NaCl on both sides and 0.6 μ M BTX on the *trans* side. As the ion concentrations on both sides of the membrane were identical, the reversal (zero-current) potential was 0 mV. Before the beginning of this record, the conductance had increased in several discrete jumps (one such jump can be seen at arrow 1). With the membrane potential held at +30 mV, both chambers were stirred (beginning at arrow 2) and 0.3 μ M STX was added to the *cis* side (arrow 3). The disturbance just before STX addition was due to insertion of the pipette into the *cis* chamber. The current record was filtered (low-pass) at 2 Hz. *b*, Dose-response curve for STX block. Same conditions as in *a* except that the STX concentration on the *cis* side was raised in increments ($E_m = +30$ mV). In this case, the data have been corrected by subtracting a 20-pA, STX-independent current (5 μ M STX *cis* and *trans*). The curve was drawn for a blocker with a dissociation constant of 5 nM. The bars indicate standard errors.

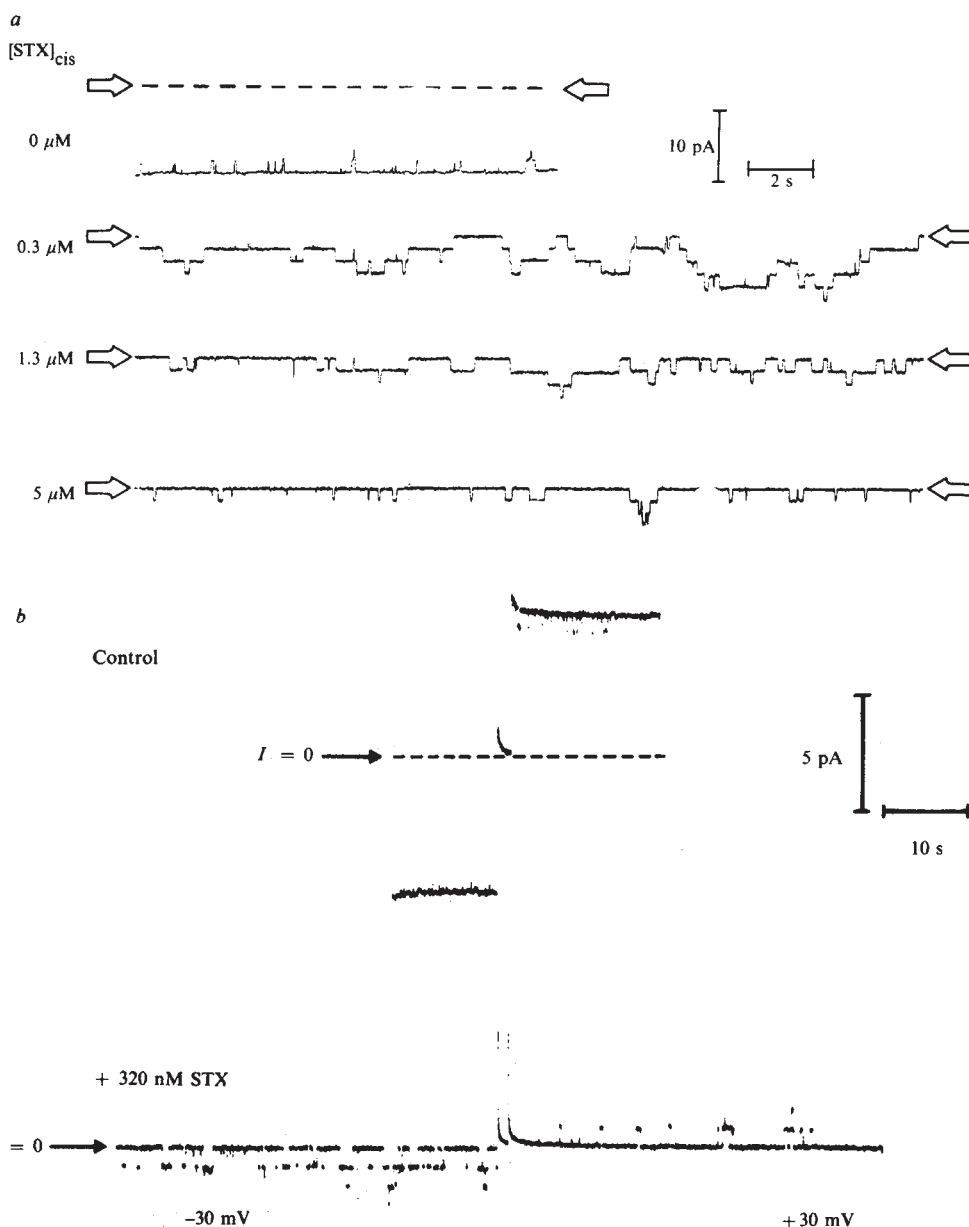


Fig. 2 *a*, Single sodium channel current fluctuations induced by STX. Sodium channels were incorporated into the bilayer as described in Fig. 1 legend. Top: control trace at -60 mV. Lower traces: same membrane at -60 mV following addition of 0.3 μ M, 1.3 μ M and 5 μ M STX to the *cis* side. The arrows indicate the zero-current level for each record. The current records were filtered (low-pass) at 20 Hz on playback. *b*, Voltage dependence of STX block. Same membrane as in *a* above before (top) and after (bottom) addition of 0.32 μ M STX to the *cis* side (voltages shown at bottom refer to both traces). BTX (0.6 μ M) was present on the *trans* side. The current records were filtered at 20 Hz on playback. The fraction of channels that remained conducting at various potentials is given below:

Potential (<i>cis-trans</i>)	% Not blocked
+60 mV	0.7
+30 mV	1.2
-30 mV	5.3
-60 mV	22.6

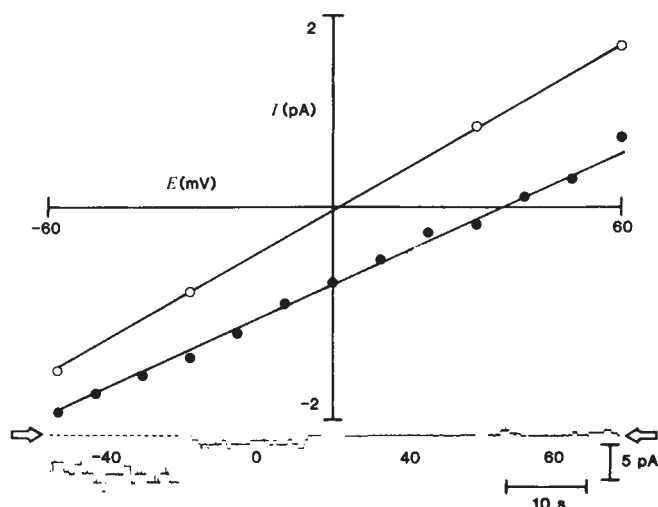


Fig. 3 Current-voltage relationships for single sodium channels. Open circles show the single-channel currents in symmetrical 0.5 M NaCl. Closed circles show the single-channel currents with 0.5 M NaCl (*trans*) and 0.1 M NaCl, 0.4 M KCl (*cis*). BTX (0.6 μ M) was present on the *trans* side. The lines are least-squares fits to the data points giving single-channel conductances of 23 pS (closed circles) and 29 pS (open circles). The insets show representative current records taken at -40, 0, +40 and +60 mV for 0.5 M NaCl (*trans*) and 0.1 M NaCl, 0.4 M KCl (*cis*). Current records in insets were filtered (low-pass) at 30 Hz.

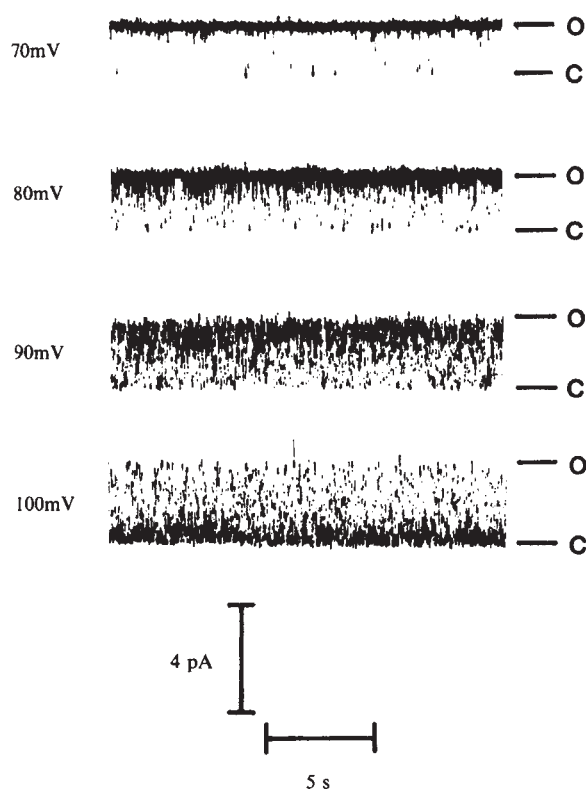


Fig. 4 Voltage dependence of BTX-activated channels. A single sodium channel was incorporated in the planar bilayer as described above. Records are shown at +70, +80, +90 and +100 mV (*cis-trans*) as indicated. 'C' and 'O' indicate the closed- and open-state current levels, respectively. The zero-current level for each record was ~ 0.3 – 0.5 pA below the closed-state level. Records were filtered at 60 Hz on playback. BTX (0.6 μ M) was present on the *trans* side. Similar results were obtained in three experiments using single-channel bilayers from two different membrane preparations.

at +60 mV (see Fig. 2 legend). Thus, voltages that would enhance entry of STX (a divalent cation) into the channel favour block. Although block of cardiac sodium channels by TTX has been reported to be voltage dependent¹⁶, this conclusion has been challenged on technical grounds¹⁷ and is not supported by more recent voltage-clamp studies^{18,19}. Such an effect has not been reported in neurones. However, because of the long relaxation times for STX block ($\tau = 1$ s at $K_d = 5$ nM; $[STX] = 100$ nM), the magnitude of the peak sodium current during a single voltage-clamp pulse reflects the degree of block at the preceding holding potential rather than at the pulse potential. Even the complex pulse protocols required to test for steady-state voltage-dependent STX block of sodium channels without BTX treatment¹⁹, can be used to explore only a limited voltage range. Our ability to examine the steady-state voltage dependence of STX block over a wide range of potentials is a consequence of the inhibition of sodium channel inactivation by BTX. We cannot exclude the possibility that the high sodium concentration or channel modification by BTX are necessary for the STX action to be voltage dependent.

Figure 3 illustrates the selectivity of single BTX-activated channels. In this experiment, the *cis* chamber was perfused with a vesicle-free solution containing 0.1 M NaCl, 0.4 M KCl. The *trans* chamber contained 0.5 M NaCl and 0.6 μ M BTX. The reversal potential for the single-channel currents in this experiment was about +36 mV ($E_{Na} = 41$ mV) which implies a permeability ratio P_K/P_{Na} of 0.06 (mean of 0.07 ± 0.02 (s.e.) in three experiments). In a similar experiment using CsCl instead of KCl, we were unable to detect any permeability to Cs (that is, the zero-current potential approximated E_{Na}). The P_K/P_{Na} ratio is lower than that obtained for BTX-modified channels in the node of Ranvier¹⁰ and in neuroblastoma cells in culture^{20,21}, and is comparable with a P_K/P_{Na} of about 0.08 for unmodified channels in the node of Ranvier^{10,22}. In general, permeability ratios depend on the ionic composition of the bathing solutions^{23,24} and can even differ significantly between two related cell lines²⁰.

The single-channel current fluctuations in symmetrical 0.5 M NaCl (Fig. 3, open circles) revealed a unit conductance of about 30 pS. This value is consistent with the reported value of about 15–18 pS using the patch-clamp technique^{2,25} (but see ref. 26), considering the much higher sodium concentration used in our experiments and assuming a saturating conductance-activity relationship²⁷. The single-channel conductance with 0.1 M NaCl and 0.4 M KCl on the *cis* side and 0.5 M NaCl on the *trans* side (Fig. 3, closed circles) was 23 pS, which was somewhat lower than in symmetrical 0.5 M NaCl solutions.

Figure 4 illustrates the voltage dependence of a single BTX-activated sodium channel incorporated into a planar bilayer. At potentials less than +70 mV, the sodium channel was open >99% of the time; at more positive potentials, the channel began to close for brief intervals, many of which were too short to be fully resolved by our amplifier. At very large potentials the channel was closed almost all the time. The midpoint of the conductance-voltage curve was about +95 mV. As the sodium channels were oriented in the bilayer with their extracellular sides (STX-blocking sites) facing *cis*, positive potentials across the bilayer corresponded to (cell interior) negative potentials *in situ*.

The specific block of the conductance by STX at nanomolar concentrations, the activation of the channels by BTX, the selectivity of the channels for sodium over potassium, and the voltage dependence together with the unit conductance, indicate that these channels are the well studied sodium channels of excitable cells. As the planar bilayer system allows control of the solutions on both sides of the membrane and the lipid composition of the bilayer, it should be useful for biochemically modifying specific sites on the channels as well as for characterizing purified sodium channels.

We thank Marc B. Blaustein and Edouard L. Noisin for excellent technical assistance, Mr G. Ay for making the chambers, Mr J. Michael for assisting with the electronics, Dr J. Daly

for a generous supply of BTX, Drs B. E. Alger, M. P. Blaustein, L. Goldman, W. J. Lederer, L. J. Mullins, D. A. Nachshen, and M. T. Nelson for comments on the manuscript. We particularly thank Drs E. Bamberg, R. Coronado, R. Latorre, C. Miller, G. Szabo and R. Waldbillig for introducing us to the mystique of the bilayer. This work was supported by NIH grants NS-16285 (B.K.K.) and NS-16775 (R.J.F.), by an A. P. Sloan Foundation fellowship (B.K.K.), and by contract DAMD-17-C-2188 from the US Army Research and Development Command (B.K.K. and R.J.F.).

Received 20 January; accepted 8 March 1983.

- Hodgkin, A. L. & Huxley, A. F. *J. Physiol., Lond.* **116**, 497–506 (1952).
- Sigworth, F. J. & Neher, E. *Nature* **287**, 447–449 (1980).
- Albuquerque, E. X. & Daly, J. in *Receptors and Recognition Ser. B*, Vol. 1 (ed. Cuatrecasas, P.) 297–338 (Chapman & Hall, London, 1976).
- Narahashi, T. *Physiol. Rev.* **54**, 813–866 (1974).
- Mueller, P., Rudin, D. O., Tien, H. T. & Wescott, W. C. *J. phys. Chem.* **67**, 534–535 (1963).
- Krueger, B. K., Ratzlaff, R. W., Strichartz, G. R. & Blaustein, M. P. *J. Membrane Biol.* **50**, 287–310 (1979).
- Miller, C. *J. Membrane Biol.* **40**, 1–23 (1978).
- Miller, C. & Racker, E. *J. Membrane Biol.* **30**, 283–300 (1976).
- Cohen, F. S., Zimmerberg, J. & Finkelstein, A. *J. gen. Physiol.* **75**, 251–270 (1980).
- Khodorov, B. I. in *Membrane Transport Processes* Vol. 2 (eds Tosteson, D. C., Ovchinnikov, Y. A. & Latorre, R.) 153–174 (Raven, New York, 1978).
- Huang, L.-Y. M., Moran, N. & Ehrenstein, G. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2082–2085 (1982).
- Wagner, H.-H. & Ulbricht, W. *Pflüger's Arch. ges. Physiol.* **359**, 297–315 (1975).
- Krueger, B. K. & Blaustein, M. P. *J. gen. Physiol.* **76**, 287–313 (1980).
- Tamkun M. M. & Catterall, W. A. *Molec. Pharmac.* **19**, 78–86 (1981).
- Ritchie, J. M. & Rogart, R. B. *Rev. Physiol. Biochem. Pharmac.* **79**, 1–50 (1977).
- Baer, M., Best, P. M. & Reuter, H. *Nature* **263**, 344–345 (1976).
- Cohen, I. S. & Strichartz, G. R. *Biophys. J.* **17**, 275–279 (1977).
- Colatsky, T. J. & Gadsby, D. C. *J. Physiol., Lond.* **306**, 20P (1980).
- Cohen, C. J., Bean, B. P., Colatsky, T. J. & Tsien, R. W. *J. gen. Physiol.* **78**, 383–411 (1981).
- Huang, L.-Y. M., Catterall, W. A. & Ehrenstein, G. *J. gen. Physiol.* **73**, 839–854 (1979).
- Frelin, C., Vigne, P. & Lazdunski, M. *Eur. J. Biochem.* **119**, 437–442 (1981).
- Hille, B. *J. gen. Physiol.* **59**, 637–658 (1972).
- Ebert, G. A. & Goldman, L. *J. gen. Physiol.* **68**, 327–340 (1976).
- Begenisich, T. B. & Cahalan, M. D. *J. Physiol., Lond.* **307**, 217–242 (1980).
- Patlak, J. & Horn, R. *J. gen. Physiol.* **79**, 333–351 (1982).
- Quandt, F. N. & Narahashi, T. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6732–6736 (1982).
- Hille, B. *J. gen. Physiol.* **66**, 535–560 (1975).

Possible association of alcohol tolerance with increased synaptic Ca^{2+} sensitivity

Marina A. Lynch & John M. Littleton

Department of Pharmacology, King's College London, Strand, London WC2R 2LS, UK

The inhibitory effect of ethanol on neurotransmitter release^{1,2} has been suggested to be due to either reduced Ca^{2+} entry³ or increased removal of free intracellular Ca^{2+} from the synapse⁴. The use of the Ca^{2+} ionophore, A23187, to allow direct access of external Ca^{2+} to the presynaptic interior⁵ should help to determine which of these two factors is the more important, as ethanol should inhibit A23187-induced release of transmitter only if increased Ca^{2+} removal from the synapse is important. Here we show in rat striatal slices that, although ^3H -dopamine release evoked by depolarization with 40 mM K^+ is inhibited by 50 mM ethanol, the release evoked by A23187 is enhanced by the presence of ethanol *in vitro*. The results suggest that ethanol reduces depolarization-induced transmitter release by reducing Ca^{2+} entry to the presynaptic terminal. However, for brain slices taken from rats made tolerant to ethanol, ^3H -dopamine release in the absence of ethanol showed altered characteristics; both K^+ depolarization and A23187 released a significantly greater fraction of ^3H -dopamine from these slices than from controls. Thus tolerance to the inhibitory effect of ethanol on release may develop by a mechanism involving increased sensitivity of the terminal to Ca^{2+} entry.

In the experiments described here we used a superfusion system in which rat striatal slices preloaded with ^3H -dopamine were subjected to two periods of stimulation (S_1 and S_2) either by K^+ depolarization or by superfusion with A23187. In most of the experiments, Ca^{2+} (2 mM) was present in the superfusate throughout, but in some (see Table 1) the superfusate was

Table 1 Effect of ethanol administration *in vivo* on the characteristics of ^3H -dopamine efflux from superfused rat striatal slices

Release induced by:	$[\text{Ca}^{2+}]_{\text{ext}}$	S	% Fractional release of ^3H min ⁻¹ ($\pm$ s.e.m.)	
			Control	Ethanol-tolerant
40 mM K^+	200 μM	S_1	0.36 ± 0.01	$0.57^* \pm 0.01$
40 mM K^+	200 μM	S_2	0.24 ± 0.01	$0.32^* \pm 0.01$
40 mM K^+	2 mM	S_1	1.69 ± 0.03	$2.55^* \pm 0.06$
40 mM K^+	2 mM	S_2	1.42 ± 0.05	$1.94^* \pm 0.06$
A23187 (12 μM)	2 mM	S_1	0.56 ± 0.05	$1.33^+ \pm 0.18$
A23187 (12 μM)	2 mM	S_2	0.54 ± 0.08	$2.29^+ \pm 0.33$

The results given are derived from the d.p.m. ^3H released per min in S_1 and S_2 . A correction is made for the ^3H -release occurring spontaneously before S_1 and S_2 and the remaining value expressed as a percentage of the total ^3H remaining in the tissue at the beginning of each collection period. This value, termed the % fractional release per min, is shown in the table (mean $\pm$ s.e.m. of at least five experiments). The mean values for S_1 and S_2 between control and ethanol-tolerant groups have been compared by Student's *t*-test. Comparisons were always made between experiments on 'test' and 'control' preparations in equal numbers carried out on the same day and randomly distributed between the six sets of superfusion apparatus. The values shown for K^+ -induced release differ from those given in the text because the experimental design is different. The superfusate is $[\text{Ca}^{2+}]$ -free throughout, except for the two periods of stimulation. $[\text{Ca}^{2+}]_{\text{ext}}$, external Ca^{2+} concentration.

* $P < 0.01$; + $P < 0.05$.

Ca^{2+} -free until the periods of stimulation. Ethanol was occasionally included in the second period of stimulation to assess its effects on ^3H -dopamine efflux *in vitro*. Further details of the superfusion system and the method of calculating results are given in the legend to Table 1.

When ^3H -dopamine release was evoked by superfusion with 40 mM K^+ the presence of 50 mM ethanol during S_2 significantly ($P < 0.01$ in paired Student's *t*-test) reduced the fractional release per min compared with that in S_1 [$\text{S}_1 = 2.66 \pm 0.25\%$; $\text{S}_2 = 2.19 \pm 0.17\%$ (mean $\pm$ s.e.m., $n = 6$)]. In the absence of ethanol there was no significant difference between mean fractional release of ^3H -dopamine per min in the two periods ($\text{S}_1 = 2.26 \pm 0.26\%$; $\text{S}_2 = 2.15 \pm 0.22\%$, $n = 6$). This small reduction by ethanol of transmitter release evoked by K^+ depolarization is in agreement with the findings of others². However, when release of ^3H -dopamine was evoked in S_1 and S_2 by A23187 (12 μM) present in the superfusate, the addition of 50 mM ethanol during S_2 did not reduce release, but enhanced it. Thus, in the absence of ethanol the mean fraction of ^3H -dopamine released per min in S_1 was $0.56 \pm 0.05\%$ and in S_2 , $0.54 \pm 0.08\%$; when ethanol was present in S_2 the respective values were $0.73 \pm 0.05\%$ and $0.96 \pm 0.08\%$, which is a significant increase ($P < 0.01$ in paired Student's *t*-test, $n = 5$ in all cases). This enhancing effect of ethanol on A23187-induced release may be due to greater mobility of the ionophore in a membrane environment made more fluid by the presence of ethanol⁶. This finding suggests that Ca^{2+} redistribution within the synapse is unlikely to be important in the inhibitory effect of ethanol on transmitter release produced by the more physiological stimuli of electrically- or K^+ -induced depolarization. Inhibition of Ca^{2+} entry through voltage-dependent channels³ is, therefore, the more likely mechanism.

If ethanol is given chronically to rodents by inhalation, tolerance to the intoxicating effect of the drug develops rapidly⁷. This tolerance has been attributed to synaptic adaptation in which synaptic membrane lipids are thought to play a part^{8,9}. Such changes might produce tolerance by excluding ethanol from a membrane site of action¹⁰ or might alter the characteristics of synaptic transmission to overcome the inhibitory effects of ethanol. Appropriate adaptations would be an alteration of the characteristics of Ca^{2+} entry through the voltage-dependent channels or an increase in the sensitivity of the presynaptic

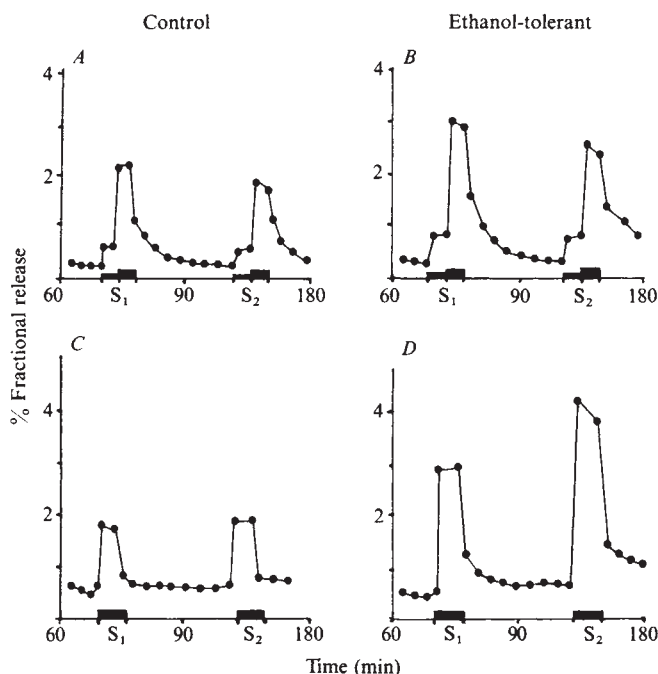


Fig. 1 Superfusion of rat striatal slices from control rats and rats made tolerant to ethanol. The mean percentage fractional release of ^3H -dopamine in each collection period is shown for slices from control (A, C) and ethanol-tolerant (B, D) rats ($n=5$ in all groups). A, B, K^+ -induced; C, D, A23187-induced. The thickened lines on the abscissa indicate S_1 and S_2 ; two phases are shown in the K^+ -induced experiments ($200\ \mu\text{M}$ and $2\ \text{mM}\ \text{Ca}^{2+}$, respectively). Twenty male Sprague-Dawley rats were divided randomly into four groups of five each. Two groups received ethanol by inhalation for 5–7 days (ethanol-tolerant rats) while the remaining two groups were kept in the same conditions but received no ethanol (controls). Rats were killed by stunning and decapitation, the brains were rapidly removed and the striatum dissected on ice. Tissue prisms ($0.2\ \text{mm} \times 0.2\ \text{mm} \times 0.2\ \text{mm}$), prepared using a McIlwain chopper, were incubated in oxygenated Krebs solution containing ^3H -dopamine (Amersham) at 37°C for 1 h. The prisms were then filtered, rinsed, placed on filter paper circles in tissue holders and superfused with oxygenated Krebs solution at 37°C for 1 h before sample collection. Tissues were stimulated at two periods, S_1 and S_2 , during the experiment. In the experiments where stimulation was induced by K^+ (A and B), Ca^{2+} -free Krebs solution was used except during S_1 and S_2 , when $40\ \text{mM}\ \text{K}^+$ was present in the superfusate which also contained $200\ \mu\text{M}\ \text{Ca}^{2+}$ for the first 5 min and $2\ \text{mM}\ \text{Ca}^{2+}$ for the second 5 min of each stimulation period. A23187-induced ^3H release was produced by superfusing slices with Krebs solution containing $12\ \mu\text{M}$ A23187 for 7 min during S_1 and S_2 ; $2\ \text{mM}\ \text{Ca}^{2+}$ was present in the superfusate throughout these experiments (C and D). Note that this affects the spontaneous release of transmitter; it is higher in experiments C and D than in A and B where Ca^{2+} -free Krebs was used as the superfusate.

This design was chosen to investigate the possibility that an altered affinity of Ca^{2+} for the Ca^{2+} channel occurs in ethanol tolerance. For example, if the affinity for Ca^{2+} had increased one would expect a greater percentage enhancement of release at the lower Ca^{2+} concentration. At both external Ca^{2+} concentrations and in S_1 and S_2 , the fraction of ^3H -dopamine released by K^+ depolarization was significantly greater in slices from ethanol-tolerant rats than in slices from controls. Similar findings have been reported for electrically evoked transmitter release¹¹. This suggests that the characteristics of neurotransmitter release have been altered, but does not distinguish between increased Ca^{2+} entry and increased Ca^{2+} sensitivity as the likely mechanism. There was no evidence for a differential effect on low or high Ca^{2+} concentration, suggesting that no alteration has occurred in the affinity for Ca^{2+} of the voltage-dependent Ca^{2+} channel.

When ^3H -dopamine release was evoked by superfusion with $12\ \mu\text{M}$ A23187, the pattern was very similar. The fraction of ^3H -dopamine released from brain slices of ethanol-tolerant rats was much greater than that from control animals (Table 1 and Fig. 1). It should be emphasized that it is extremely unlikely that pharmacological concentrations of ethanol remained in the tissue at this time. Apart from the time spent in preparing and incubating the brain slices, they were subjected to ethanol-free superfusion for at least 1 h before presentation of the stimulus for release in S_1 . This increased release of neurotransmitter can therefore be regarded as an effect of withdrawal of ethanol and, as such, could provide the basis for both tolerance and dependence.

The increase in release of dopamine induced by the ionophore is unlikely to be simply a consequence of the previously reported altered synaptic membrane lipid composition^{8,9} as these alterations would tend to make the membrane less fluid and thus reduce the effects of the ionophore⁶. Also, although other interpretations, such as reduced endogenous levels or synthesis of dopamine in the tissues of ethanol-tolerant rats are possible, these do not accord with previously published results¹². In our view the present findings strongly suggest that the increased fractional release of dopamine observed is a consequence of an increased sensitivity of the presynaptic terminal to Ca^{2+} entry in brain slices from ethanol-tolerant rats. Experiments on the relation between external Ca^{2+} concentration and K^+ depolarization-induced release and on the relation between A23187 concentration and release (not shown) suggest that a 3–4-fold increase in Ca^{2+} sensitivity is necessary to explain the magnitude of the increase.

This proposed change in sensitivity of the presynaptic terminal to Ca^{2+} entry may have an important role in the molecular basis of tolerance to ethanol. It may also persist into the ethanol physical withdrawal syndrome (our unpublished preliminary experiments). The changes in synaptic membrane lipid composition which we believe may underlie this altered sensitivity have also been reported recently in the striatum of rats made tolerant to morphine¹³. This mechanism may therefore be of broader significance in the development of tolerance to central depressant drugs.

We thank Dr J Halliday for helpful comments. This research was supported by grants from the MRC and the Medical Council on Alcoholism Ltd.

Received 5 November 1982; accepted 23 March 1983.

1. Carmichael, F. J. & Israel, Y. *J. Pharmac. exp. Ther.* **193**, 824–834 (1975).
2. Kalant, H. & Grose, W. J. *Pharmac. exp. Ther.* **158**, 386–393 (1967).
3. Harris, R. A. & Hood, W. F. J. *Pharmac. exp. Ther.* **213**, 562–568 (1980).
4. Michaelis, E. K. & Myers, S. L. *Biochem. Pharmac.* **28**, 2081–2087 (1979).
5. Akerman, K. E. O. & Nicholls, D. J. *Eur. J. Biochem.* **115**, 67–73 (1981).
6. Deleers, M., Couturier, E. & Malaisse, W. J. *Cell Calcium* **2**, 159–171 (1981).
7. Grieve, S. J. & Littleton, J. M. J. *Pharm. Pharmac.* **31**, 605–610 (1979).
8. Littleton, J. M., John, G. R. & Grieve, S. J. *Alcohol. clin. exp. Res.* **3**, 50–56 (1979).
9. Chin, J. H., Parsons, L. M. & Goldstein, D. B. *Biochim. biophys. Acta* **513**, 358–363 (1978).
10. Rottenberg, H., Waring, A. & Rubin, E. *Science* **213**, 583–584 (1981).
11. Clark, J. W., Kalant, H. & Carmichael, F. J. *Can. J. Physiol. Pharmac.* **55**, 758–768 (1977).
12. Tabakoff, B. in *Alcohol and Opiates* (ed. Blum, K.) 21–40 (Academic, New York, 1977).
13. Heron, D. S., Shinitzky, M., Zamir, N. & Samuel, D. *Biochem. Pharmac.* **31**, 2435–2438 (1982).

terminal to the reduced Ca^{2+} entry occurring in the presence of ethanol.

To test these alternative hypotheses, we first made a group of rats tolerant to ethanol by exposing them to ethanol at concentrations of $10\text{--}20\ \text{mg l}^{-1}$ in inspired air for 5–7 days. Rats were severely intoxicated on several occasions during this regime and when they were killed, all had blood ethanol concentrations (gas-liquid chromatography of mixed arteriovenous blood from the neck) in excess of $50\ \text{mM}$ ($230\ \text{mg per } 100\ \text{ml}$). Corpus striatum slices of brain from these animals were compared with those of sham-treated controls. Further details are given in the legend to Table 1.

The slices of striatum from the two groups showed no difference in their ability to take up ^3H -dopamine (not shown). The release of ^3H -dopamine evoked by $40\ \text{mM}\ \text{K}^+$ was studied at two external Ca^{2+} concentrations (see Table 1 and Fig. 1).

A nucleotide regulatory site for somatostatin inhibition of adenylate cyclase in S49 lymphoma cells

Karl H. Jakobs, Klaus Aktories & Günter Schultz

Pharmakologisches Institut der Universität Heidelberg,
Im Neuenheimer Feld 366, D-6900 Heidelberg, FRG

The cyc^- variants of S49 lymphoma cells have served as powerful tools for studying the components and mechanisms of hormone-induced adenylate cyclase stimulation, as these cells are deficient in the guanine nucleotide regulatory site (N_s) mediating hormone, guanine nucleotide, cholera toxin and fluoride-induced stimulations of the enzyme¹. Because of this deficiency, membranes of these cells have been used for reconstitution of the system by inserting the coupling component derived from other cell types²⁻⁴. The hormone-sensitive adenylate cyclase is not only stimulated by hormones but can also be inhibited by a wide variety of hormones and neurotransmitters⁵⁻⁸, and there is some evidence that hormonal inhibition may be mediated by a distinct guanine nucleotide regulatory site⁵⁻⁸. Studies in cyc^- cells lacking a functional N_s may therefore answer this unresolved, important question. We have recently observed⁹ that stable GTP analogues can inhibit cyc^- adenylate cyclase stimulated by purified, preactivated N_s or forskolin, which can activate adenylate cyclase even in the absence of a functional N_s (ref. 10). The data indicated that these N_s -deficient cells contain an inhibitory guanine nucleotide site, N_i . To strengthen this concept, we investigated whether the cyc^- adenylate cyclase can be inhibited by a hormone. We report here that somatostatin decreases cyclic AMP levels in cyc^- cells, inhibits the forskolin-stimulated adenylate cyclase and causes a concomitant increase in a high affinity GTPase activity in cyc^- membranes. The data strongly suggest that both the hormone- and guanine nucleotide-induced adenylate cyclase inhibitions in cyc^- cells are mediated by N_i and that the mechanisms of activation and inactivation of N_i are similar to those established for N_s .

In the presence of forskolin (30 μ M) and the cyclic nucleotide phosphodiesterase inhibitor, 3-isobutyl-1-methylxanthine (30 μ M), somatostatin decreased the cyclic AMP levels in cyc^- cells by up to 35% (Table 1). Half-maximal and maximal attenuations were observed at about 10^{-10} and 10^{-8} M somatostatin, respectively. In cyc^- membranes, forskolin caused a concentration-dependent increase in adenylate cyclase activity, which was not maximal even at 100 μ M (Fig. 1), as reported previously¹⁰. Forskolin activation of the enzyme was not attenuated by somatostatin. However, in the presence of GTP (3 μ M), which by itself decreased the adenylate cyclase activity by ~20%, somatostatin (1 μ M) caused an additional decrease in activity of 20–35%, which was observed at all forskolin concentrations studied (1–100 μ M). Half-maximal and maximal effects of somatostatin in cyc^- membranes were observed at about 1 and 30 nM, respectively, and were not mimicked by various other hormones and neurotransmitters studied, including angiotensin II, [Met]enkephalin, α -adrenergic or muscarinic cholinergic agonists (data not shown).

The stable GTP analogue, guanosine 5'-O-(3-thiotriphosphate) (GTP γ S), inhibited the forskolin-stimulated cyc^- adenylate cyclase with a half-maximal effect at ~5 nM (Fig. 2)⁹. The inhibition was time-dependent and was not abolished by washing the membranes⁹. Somatostatin (1 μ M) had no effect on control activity but apparently increased the inhibitory potency of GTP γ S by a factor of 2–2.5. This apparent increase in GTP γ S-induced inhibition was at least partially due to a somatostatin-induced acceleration of the inhibitory action of GTP γ S (not shown).

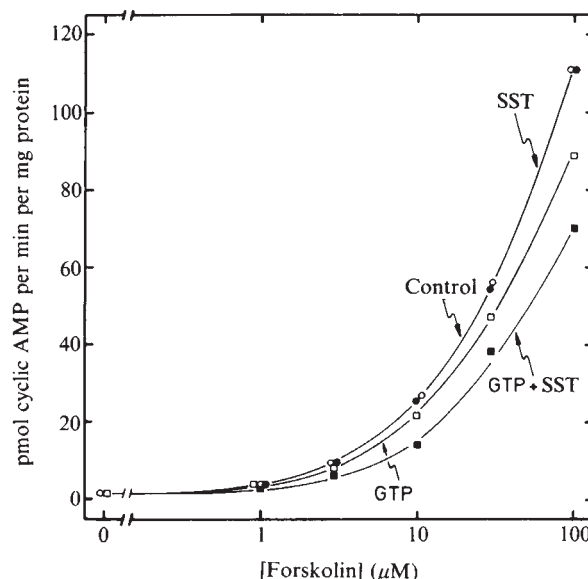


Fig. 1 Influence of somatostatin on forskolin-stimulated adenylate cyclase in S49 lymphoma cyc^- membranes. Adenylate cyclase activity was determined essentially as described⁹ with a reaction mixture containing 50 μ M [α -³²P]ATP (0.4 μ Ci per tube), 100 μ M MnCl₂, 1 mM 3-isobutyl-1-methylxanthine, 0.1 mM cyclic AMP, 5 mM creatine phosphate, 0.4 mg ml⁻¹ creatine kinase, 1 mg ml⁻¹ bacitracin, 2 mg ml⁻¹ bovine serum albumin and 50 mM triethanolamine-HCl, pH 7.4, without (Control) and with 1 μ M somatostatin (SST), 3 μ M GTP or GTP plus somatostatin at the indicated concentrations of forskolin. The reactions were initiated by the addition of cyc^- membranes (20–30 μ g protein per tube) prepared as described before⁹ and conducted for 5 min at 37 °C. Isolation of cyclic AMP formed was carried out as previously described²⁰. Mean values of triplicates are given which did not differ more than 5% of the means.

Table 1 Influence of somatostatin on cyclic AMP levels in S49 cyc^- lymphoma cells

	pmol cyclic AMP per 10 ⁶ cells
Control	3.20 ± 0.17
Somatostatin 10 ⁻¹⁰ M	2.46 ± 0.12
10 ⁻⁹ M	2.12 ± 0.08
10 ⁻⁸ M	2.06 ± 0.15
10 ⁻⁷ M	2.03 ± 0.09

S49 lymphoma cyc^- variants (6.8×10^6 cells per tube) were pretreated for 10 min at 36 °C with 30 μ M forskolin and 30 μ M 3-isobutyl-1-methylxanthine in serum-free Dulbecco's modified Eagle's medium containing 10 mM HEPES, pH 7.4. Thereafter, somatostatin at various concentrations was added and the cells were incubated for a further 5 min. The incubation was terminated by the addition of trichloroacetic acid (10% final concentration) and ³H-labelled cyclic AMP was added to monitor the recovery during the isolation procedure. Cyclic AMP was isolated on acid alumina columns as described elsewhere¹⁸ with the modification that before elution of cyclic AMP the columns were washed with 5 ml of a solution containing 0.6 M HCl in 50% ethanol. Cyclic AMP was eluted with 4 ml of 0.2 M acetic acid. After lyophilization, the samples were dissolved in 0.2 M sodium acetate, pH 6.5, and in aliquots cyclic AMP content (by a radioimmunoassay¹⁹) and individual cyclic AMP recoveries (50–70%) were determined. Values given are the mean of cyclic AMP ± s.e.m. ($n = 8$) per 10⁶ cells.

Inactivation of the hormone-activated stimulatory coupling component, N_s , is thought to be caused by the hydrolysis of N_s -bound GTP to GDP and P_i (ref. 11). If the inhibitory regulatory site, N_i , is inactivated by similar mechanisms to N_s , somatostatin should increase GTP hydrolysis in cyc^- membranes. Figure 3 shows the effect of somatostatin on the hydrolysis of [γ -³²P]GTP in cyc^- membranes measured at various concentrations of unlabelled GTP. As observed in many other cell types¹¹⁻¹⁴, there were at least two GTPase activities in cyc^- membranes, one with high affinity for GTP ($K_m \approx 0.3 \mu$ M) and one with low affinity ($K_m \approx 50 \mu$ M). Somatostatin (1 μ M)

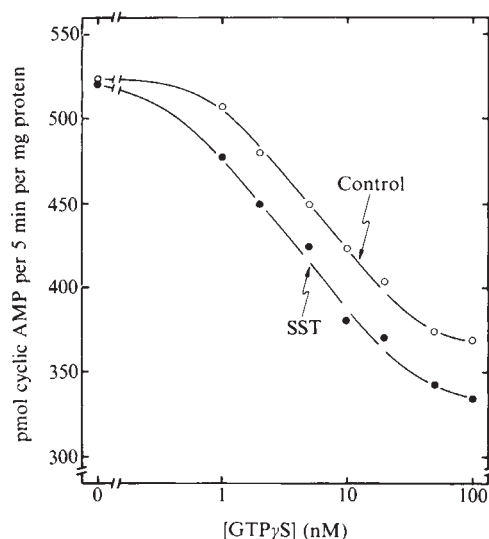


Fig. 2 Influence of somatostatin on GTP γ S-induced inhibition of cyc⁻ adenylate cyclase. In the presence of 100 μ M forskolin, the influence of somatostatin (SST, 1 μ M) was studied on cyclic AMP accumulation in cyc⁻ membranes, measured at the indicated concentrations of GTP γ S. Because the rate of cyclic AMP formation is not linear in this condition⁹, cyclic AMP accumulated over the 5-min incubation period per mg protein is given.

increased the activity of the high affinity GTPase by 30–40% but had no effect on the low affinity GTPase. The somatostatin-induced stimulation of GTP hydrolysis by the high affinity GTPase in cyc⁻ membranes was half-maximal and maximal at about 0.3 and 1 μ M GTP, respectively. Isoprenaline, a β -adrenergic agonist which stimulates adenylate cyclase in wild-type S49 lymphoma cells¹, had no effect on the basal or forskolin-stimulated cyc⁻ adenylate cyclase activities, nor on GTP hydrolysis in cyc⁻ membranes (not shown).

Thus, in S49 cyc⁻ lymphoma cells, which lack a functional N_s ¹, the adenylate cyclase is apparently inhibited via a separate guanine nucleotide regulatory site, N_i . In membranes of wild-type cells, somatostatin caused a similar inhibition of the adenylate cyclase (manuscript in preparation) to that observed in cyc⁻ membranes, thus ruling out the possibility that the cyc⁻ adenylate cyclase inhibition is due to an abnormal N_s in these membranes. The data obtained with stable GTP analogues⁹ and those reported here with somatostatin strongly suggest that the inhibitory guanine nucleotide site of the system is activated and inactivated by similar mechanisms to the stimulatory site missing in cyc⁻ cells. Activation of N_i is apparently brought about by the binding of GTP (or a stable GTP analogue). This turn-on reaction is amplified by an inhibitory hormone–receptor interaction. Inactivation of N_i is apparently caused by a high affinity GTPase activity, which may be an integral part of N_i , as also suggested but not proven for N_s (refs 11, 15). The GTPase activity associated with N_i , however, is not affected by cholera toxin, which inhibits the N_s -associated GTPase stimulation^{16,17}. Furthermore, cholera toxin did not amplify the inhibitory potency or efficiency of GTP in cyc⁻ membranes⁹, whereas the toxin is known to amplify the adenylate cyclase stimulation by GTP^{15,16}. The data do not show whether GDP is bound to N_i in the inactive state. Activation of N_i can also be caused by the binding of stable GTP analogues such as GTP γ S or guanylyl 5'-yl-imidodiphosphate. This activation of N_i , which occurs after a short lag period, can be competitively inhibited by GTP or by the stable GDP analogue, guanosine 5'-O-(2-thiodiphosphate)⁹, and as shown here can be accelerated and/or amplified by an inhibitory hormone–receptor interaction. Because of their resistance to the inactivating GTPase, activation of N_i and subsequent inhibition of the adenylate cyclase by the stable GTP analogues are persistent processes⁹.

Data similar to those given here for the action of an inhibitory hormone and guanine nucleotides on the adenylate cyclase were

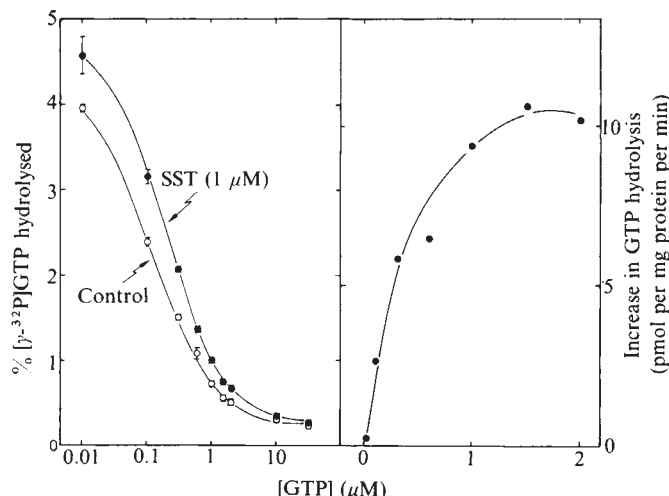


Fig. 3 Influence of somatostatin on the hydrolysis of GTP in cyc⁻ membranes. Hydrolysis of [γ -³²P]GTP (0.2 μ Ci per tube) was determined essentially as described elsewhere¹⁴ with a reaction mixture containing 1 mM 3-isobutyl-1-methylxanthine, 1 mM EDTA, 0.1 mM ATP, 1 mM adenylyl 5'-yl-imidodiphosphate, 3 mM MgCl₂, 5 mM creatine phosphate, 0.4 mg ml⁻¹ creatine kinase, 1 mg ml⁻¹ bacitracin, 2 mg ml⁻¹ bovine serum albumin and 50 mM triethanolamine-HCl, pH 7.4, without and with 1 μ M somatostatin (SST) at the indicated concentrations of unlabelled GTP. The reactions were initiated by the addition of cyc⁻ membranes (5 μ g protein per tube) and conducted for 10 min at 25°C. Isolation of ³²P_i formed was carried out as described¹⁴. Shown on the left panel is the isotope dilution curve of [γ -³²P]GTP hydrolysis (%) measured at various concentrations of GTP (log scale). Mean values $\pm$ s.e.m. ($n = 3$) are given. The right panel shows the increase in GTPase activity induced by somatostatin at various concentrations of GTP (linear scale).

obtained in and reported for many other cell types. However, as normal cells apparently contain both the stimulatory and the inhibitory guanine nucleotide sites, it was not readily possible to distinguish between the actions of guanine nucleotides on these two coupling components. The S49 cyc⁻ lymphoma cells, which are deficient in N_s (ref. 1), are thus not only useful probes for understanding the stimulatory coupling of hormone receptors to adenylate cyclase but, as shown here, may be even more useful for studying the inhibitory regulation of the adenylate cyclase by hormones and guanine nucleotides.

We thank Ms Gabriele Gabel and Mrs Christina Stannek for technical assistance and Dr W. Roth, Serono, and Dr H. Metzger, Hoechst AG, for donating somatostatin and forskolin, respectively. This work was supported by the Deutsche Forschungsgemeinschaft.

Received 22 December 1982; accepted 16 March 1983.

- Johnson, G. L., Kaslow, H. R., Farfel, Z. & Bourne, H. R. *Adv. Cyclic Nucleotide Res.* **13**, 1–37 (1980).
- Northup, J. K. et al. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6516–6520 (1980).
- Sternweis, P. C., Northup, J. K., Smigel, M. D. & Gilman, A. G. *J. biol. Chem.* **256**, 11517–11526 (1981).
- Hanski, E., Sternweis, P. C., Northup, J. K., Dromerick, A. W. & Gilman, A. G. *J. biol. Chem.* **256**, 12911–12919 (1981).
- Jakobs, K. H. *Molec. cell. Endocr.* **16**, 147–156 (1979).
- Rodbell, M. *Nature* **284**, 17–22 (1980).
- Cooper, D. M. F. *FEBS Lett.* **138**, 157–163 (1982).
- Jakobs, K. H., Aktories, K. & Schultz, G. *J. Receptor Res.* (in the press).
- Jakobs, K. H., Gehring, U., Gaugler, T., Pfeuffer, T. & Schultz, G. *Eur. J. Biochem.* **130**, 605–611 (1983).
- Seamon, K. B. & Daly, J. W. *J. Cyclic Nucleotide Res.* **7**, 201–224 (1981).
- Cassel, D. & Selinger, Z. *Biochim. biophys. Acta* **452**, 538–551 (1976).
- Koski, G. & Klee, W. A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4185–4189 (1981).
- Aktories, K. & Jakobs, K. H. *FEBS Lett.* **130**, 235–238 (1981).
- Aktories, K., Schultz, G. & Jakobs, K. H. *Molec. Pharmacol.* **21**, 336–342 (1982).
- Ross E. M. & Gilman, A. G. *Rev. Biochem. Sci.* **49**, 533–564 (1980).
- Cassel, D. & Selinger, Z. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3307–3311 (1977).
- Aktories, K., Schultz, G. & Jakobs, K. H. *FEBS Lett.* **146**, 65–68 (1982); *Biochim. biophys. Acta* **719**, 58–64 (1982).
- Jakobs, K. H., Böhme, E. & Schultz, G. in *Eukaryotic Cell Function and Growth. Regulation by Intracellular Cyclic Nucleotides* (eds Dumont, J. E., Brown, B. L. & Marshall, N.J.) 295–311 (Plenum, New York, 1976).
- Harper, J. F. & Brooker, G. *J. Cyclic Nucleotide Res.* **1**, 207–218 (1975).
- Jakobs, K. H., Saur, W. & Schultz, G. *J. Cyclic Nucleotide Res.* **2**, 381–392 (1976).

A binary plant vector strategy based on separation of *vir*- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid

A. Hoekema, P. R. Hirsch*, P. J. J. Hooykaas & R. A. Schilperoort

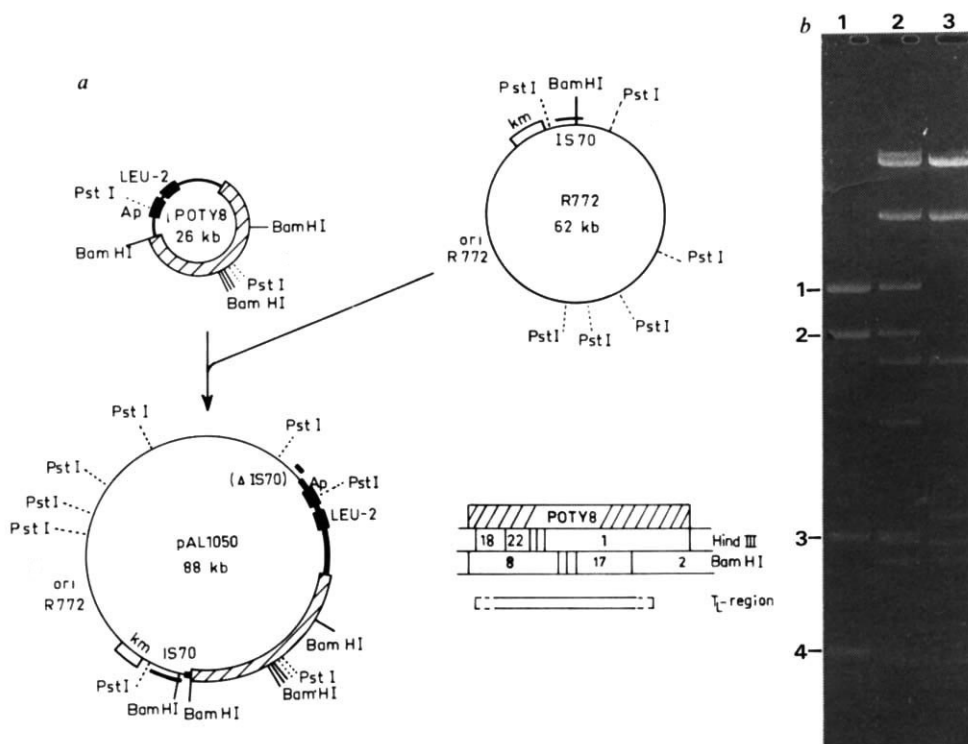
Department of Biochemistry, University of Leiden, Wassenaarseweg 64, 2333 AL Leiden, The Netherlands

The soil bacterium *Agrobacterium tumefaciens* is a plant pathogen that causes crown-gall tumours after infection of wounded dicotyledonous plants. Large plasmids (Ti-plasmids) are responsible for the oncogenicity of the bacterium^{1,3}. Crown-gall tumours contain a DNA segment, called the T-DNA, which is homologous with a defined part of the Ti-plasmid present in the tumour-inducing bacterium, and is stably integrated into the plant genome⁴⁻⁷. Apart from the T-DNA another region of the Ti-plasmid-called the *vir*-region, is essential for tumour induction⁸⁻¹¹. We report here the interaction of two compatible plasmids, one containing the *vir*-region, the other carrying the T-DNA on a wide host-range replicon. An *A. tumefaciens* strain harbouring both plasmids has a normal tumour-inducing capacity, although neither plasmid is functional alone. With this approach, the T-DNA on one plasmid can, because of its size, be easily genetically manipulated using *Escherichia coli* as a host. Transfer of this plasmid into an *A. tumefaciens* strain harbouring the plasmid with the *vir*-region allows introduction of the manipulated T-DNA into plant cells. In this way, sophisticated binary vector systems for plant genetic engineering can be developed.

Fig. 1 a, The construction of pAL1050. The part of the T-DNA of pTiAch5 present on pOTY8 is shown in the bottom right-hand corner. pOTY8 was constructed from subclones of *Bam*HI fragment 8 and *Hind*III fragment 1, using pJDB 207 as a vector plasmid. For pAL1050, R772 and pOTY8, the restriction sites for *Bam*HI and *Pst*I are indicated. Analysis with many other restriction endonucleases revealed the presence of all T-DNA fragments of pOTY8 (data not shown). For R772 a physical map has been constructed with many enzymes¹⁸. As shown above, the insertion sequence present on R772, IS70, mediated co-integrate formation between pOTY8 and R772. At both junctions of R772 and pOTY8 a copy of IS70 is present. Co-integrates formed between Ti-plasmids and R-plasmids usually dissociate after transfer from *A. tumefaciens* to *E. coli*¹⁹. This is due to recombination between the two IS elements bordering the fusion sites. Deletion of one of the copies of the IS generates stable co-integrates¹⁸⁻²⁰. The co-integrate pAL1050 was shown to be perfectly stable in our experiments, which is explained by the fact that one of the copies of IS70 was shown to be deleted (data not shown). **b**, An agarose gel showing the *Bam*HI+*Pst*I digest patterns of pOTY8 (lane 1), pAL1050 (lane 2) and R772 (lane 3). R772 has inserted into fragment 4 of pOTY8 to generate pAL1050. Clearly, fragment 1 of pOTY8, containing the *LEU2* locus itself, is unaltered, while the expression of *LEU2* is blocked on pAL1050. Presumably this inactivation is due to the polar effect of the integration of R772. The T-DNA part of this plasmid, however, is still intact.

To obtain a plasmid that contains the entire T-DNA of the octopine Ti-plasmid pTiAch5 and that can replicate autonomously in *A. tumefaciens*, as well as in *E. coli*, we used the recombinant plasmid pOTY8. pOTY8 was constructed by inserting the T-DNA of pTiAch5 into the tetracycline resistance locus of pJDB207¹² (see Fig. 1), and also carries the ampicillin resistance gene of pAT153¹³ and a *LEU2* gene as genetic markers. Details of the construction of pOTY8 will be described elsewhere (P. R. Hirsch, manuscript in preparation). As pOTY8 cannot replicate in *A. tumefaciens*, we converted it into a wide host-range plasmid by fusion with the IncP plasmid R772. To this end, R772 was introduced into *E. coli* strain HB101 (pOTY8) by conjugation, and transconjugants were used as donors in further crosses with the *A. tumefaciens* strain LBA202. Transconjugants of these crosses inheriting the ampicillin resistance marker of pOTY8 were expected to be carrying a co-integrate plasmid of pOTY8 and R772, as pOTY8 itself is non-conjugative and cannot replicate in *Agrobacterium*. The site of co-integration of R772 could be in either the vector or the T-DNA region of pOTY8.

To isolate a co-integrate with an intact T-DNA, we screened for a plasmid whose site of integration was the *LEU2* locus. Thirty of the co-integrate-carrying transconjugants were conjugated with the *E. coli* strain JA221¹⁴ (C600 *trpE leuB*) and the offspring was tested for leucine auxotrophy. One of the 30 transconjugant strains failed to grow on a minimal medium without added leucine, and so presumably harboured an R772::pOTY8 co-integrate plasmid, whose site of co-integration has led to inactivation of the *LEU2* gene. Restriction endonuclease digestion patterns of this R772::pOTY8 plasmid (called pAL1050) were analysed in detail. Figure 1 presents evidence that plasmid pAL1050 had an insertion of R772 into the pJDB207 region of pOTY8, leaving the T-DNA region unaltered. The structure of pAL1050 was further analysed by Southern hybridizations¹⁵, using labelled plasmid DNAs of R772 and pOTY8 as probes (data not shown).



* Present address: Cancer Research Campaign, Department of Biochemistry, Imperial College, London SW7 2AZ, UK.

Table 1 Plant tumour-induction tests

Strain*	Plasmids	Tomato		Kalanchoë		Tobacco		Pea	
		Tumour	OCS†	Tumour	OCS	Tumour	OCS	Tumour	OCS
LBA 4001	Cr, pTiAch5	+	+	+	+	+	+	+	+
LBA 4404	Cr, pAL 4404	—	—	—	—	—	—	—	—
LBA 1050	Cr, pAL 1050	—	—	—	—	—	—	—	—
LBA 4434	Cr, pAL 1050, pAL 4404	+	+	+	+	+	+	+	+

* The strains LBA 4001, 4404 and 4434 have Ach5 as chromosomal background. LBA 1050 has a C58-C9 chromosomal background. In all these strains a large cryptic plasmid (Cr) is present.

† OCS: synthesis of octopine in the tumour, detected as described by Otten and Schilperoort²³.

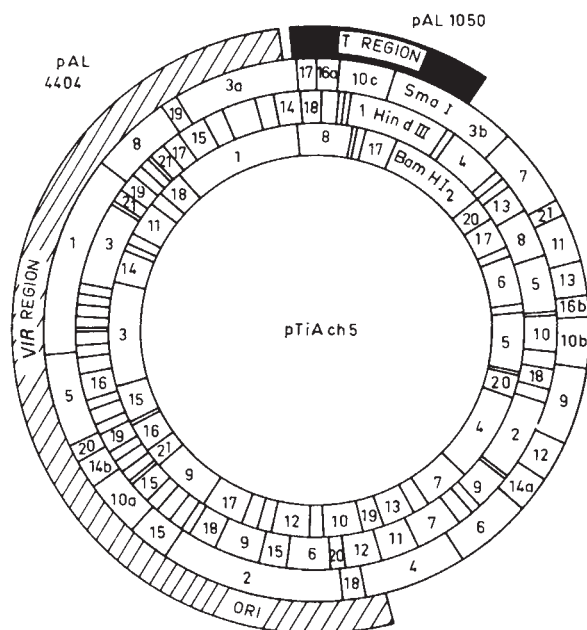


Fig. 2 The physical map of pTiAch5 is shown for the restriction endonucleases *Sma*I²¹, *Bam*HI and *Hind*III²². In the hatched and shaded areas the Ti-plasmid segments are indicated that are present separately on pAL4404 and pAL1050.

The inability of pAL1050 to restore the tumour-inducing activity of a non-oncogenic *A. tumefaciens* (which lacks an endogenous Ti-plasmid) showed that the T-DNA region alone is insufficient for tumorigenicity (Table 1).

Subsequently, pAL1050 was conjugatively transferred into a different non-oncogenic, *A. tumefaciens* strain, LBA 4404¹⁶ which contains a severely deleted Ti-plasmid, lacking the entire T-DNA region, but still with an intact *vir*-region (see Fig. 2). The tumorigenicity of the transconjugant strain LBA4434, carrying both the T-DNA-containing plasmid pAL1050 and the *vir*-region-containing plasmid pAL4404 as separate entities, was tested on several plant species. Strain LBA4434 induced normal tumours on all plants tested, and octopine was detected in these tumours (Table 1). These experiments show that the *vir*-region and the T-DNA of the octopine Ti-plasmid can be physically separated on different plasmids without affecting tumorigenicity of the bacterium. As no tumours were induced by *A. tumefaciens* carrying pAL1050 alone, genes on the *vir*-region must act in *trans* to effect transfer of the T-DNA to the plant cell.

One could argue that the tumorigenicity of the *A. tumefaciens* strain LBA4434 results from the formation of a co-integrate between pAL4404 and pAL1050 in a small part of the bacterial population. However, this is unlikely, for the following reasons. First, by Southern hybridization we have established that no homology exists between the two plasmids (data not shown), which excludes co-integrate formation by homologous recombination. Second, to study the possibility that co-integrates were formed by illegitimate recombination we crossed LBA4434

(pAL1050, pAL4404) with LBA4078, an erythromycin-resistant *A. tumefaciens* strain (no endogenous Ti-plasmid) as a recipient. Co-transfer of the non-conjugative plasmid pAL4404 with the IncP plasmid pAL1050 was not detected (frequency lower than 10^{-4}). Any co-integration between pAL1050 and pAL4404 must therefore be at a very low frequency, and so is unlikely to significantly contribute to tumour induction. In this regard, it should be kept in mind that infections of wound sites with mixtures of oncogenic and non-oncogenic *A. tumefaciens* strains in low ratios does not lead to tumour formation¹⁷. To verify this effect in our own tumour induction tests, we carried out a model experiment. We tested different mixtures of oncogenic and non-oncogenic bacteria, varying from a ratio of 1:1 up to 1:10⁴. We found that in our system a ratio of 1:10² resulted in a strongly attenuated tumour response, and at higher ratios tumour development was negligible. However, the size of the tumours induced by LBA4434 was the same as those induced by the wild-type strain Ach5. Thus it is very unlikely that tumour induction by LBA4434 is caused by a mixed population consisting mainly of non-oncogenic cells and a very limited number of cells containing a co-integrate plasmid.

Our results provide a new method for the introduction of genetic information into plant cells. One is no longer required to use an intact Ti-plasmid as a plant vector. A broad host-range plasmid containing T-DNA can be manipulated in *E. coli*, and subsequent introduction of such a plasmid into an *A. tumefaciens* strain already carrying a *vir*-region will bring about the transfer and integration of genetic material into plant cells.

This work was supported in part by the Netherlands Foundation of Chemical Research (SON) with financial aid from the Netherlands Organization for Advancement of Pure Scientific Research (ZWO). P.R.H. was sponsored by an EMBO short-term fellowship and the Cancer Research Campaign. We thank Drs J. Hille, R. J. M. van Veen and M. J. J. van Haaren for their critical reading of the manuscript, and Tonny Tuink for technical assistance during part of these studies.

Received 5 January; accepted 9 March 1983.

1. Van Larebeke, N. *et al. Nature* **252**, 169-170 (1974).
2. Zaenen, I., Van Larebeke, N., Teuchy, H., Van Montagu, M. & Schell, J. *J. molec. Biol.* **86**, 109-127 (1974).
3. Watson, B., Currier, T. C., Gordon, M. P., Chilton, M. D. & Nester, E. W. *J. Bact.* **123**, 255-264 (1975).
4. Chilton, M. D. *et al. Cell* **11**, 263-271 (1977).
5. Thomashow, M. F., Nutter, R., Montoya, A. L., Gordon, M. P. & Nester, E. W. *Cell* **19**, 729-739 (1980).
6. De Beuckeleer, M. *et al. Molec. gen. Genet.* **183**, 283-288 (1981).
7. Ooms, G. *et al. Cell* **30**, 589-597 (1982).
8. Ooms, G., Poulsen, J. A. & Schilperoort, R. A. *J. Bact.* **144**, 82-91 (1980).
9. Garfinkel, D. J. & Nester, E. W. *J. Bact.* **144**, 732-743 (1980).
10. De Greve, H. *et al. Plasmid* **6**, 235-248 (1981).
11. Hille, J., Klases, I. & Schilperoort, R. A. *Plasmid* **7**, 107-118 (1982).
12. Beggs, J. D. *Alfred Benson Symp.* **16**, 383-389 (1981).
13. Twigg, A. J. & Sheratt, D. *Nature* **283**, 216-218 (1980).
14. Beggs, J. D. *Nature* **275**, 104-109 (1978).
15. Southern, E. J. *J. molec. Biol.* **98**, 503-518 (1975).
16. Ooms, G., Hooykaas, P. J. J., Moolenaar, G. & Schilperoort, R. A. *Gene* **14**, 33-50 (1981).
17. Lippincott, B. B. & Lippincott, J. A. *J. Bact.* **97**, 620-628 (1969).
18. Hille, J., Van Kan, J., Klases, I. & Schilperoort, R. A. *J. Bact.* (in the press).
19. De Picker, A., De Block, M., Inzé, D., Van Montagu, M. & Schell, J. *Gene* **10**, 329-338 (1980).
20. Hooykaas, P. J. J., Den Dulk-Ras, H. & Schilperoort, R. A. *Plasmid* **4**, 64-75 (1980).
21. Chilton, M. D. *et al. Plasmid* **1**, 254-259 (1978).
22. De Vos, G., De Beuckeleer, M., Van Montagu, M. & Schell, J. *Plasmid* **6**, 249-253 (1981).
23. Otten, L. A. B. M. & Schilperoort, R. A. *Biochim. biophys. Acta* **527**, 497-500 (1978).

Gene deletions in patients with haemophilia B and anti-factor IX antibodies

F. Giannelli*†, K. H. Choo†, D. J. G. Rees†, Y. Boyd‡, C. R. Rizza§ & G. G. Brownlee†

† Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, UK

‡ Genetics Laboratory, Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK

§ Haemophilia Centre, Churchill Hospital, Oxford OX3 7LJ, UK

Christmas disease, or haemophilia B, is an inherited X-linked haemorrhagic disease which at present occurs in 798 known cases in the United Kingdom, corresponding to a frequency of about 1 in 30,000 males. Patients are deficient in the intrinsic clotting factor IX and are treated by replacement of this protein prepared from pooled plasma obtained from normal individuals. Occasionally treatment is complicated by the appearance of specific anti-factor IX antibodies. It seemed to us that this might be due to the absence of 'self' factor IX causing the immune system to regard the infused normal factor IX as foreign. The absence of all or part of the factor IX gene was an obvious possible reason for this, which we have now tested using our previously isolated gene probe¹. We have found four patients with gross gene defects.

In the United Kingdom, antibodies to factor IX are found in less than 1% of all patients with Christmas disease and in about 2.5% of patients with the severe form of the disease². Six patients with anti-factor IX antibodies are known in the United Kingdom, two of whom are related. The DNA from five of these patients was analysed by Southern blotting³ using four different genomic probes (I–IV) derived from the recombinant bacteriophage λ HIX1b (ref. 1) after some further characterization (Fig. 1 and legend). Two of these probes (I and IV) are entirely within introns while the others (II and III) contain different exons. These probes cover only a portion of the total gene, so a human factor IX cDNA probe was also used (probe V). This contains about 2,000 base pairs covering all the signal and coding regions and a major part of the 3' noncoding segment of the factor IX mRNA (D. S. Anson, J. A. Huddleston and G.G.B., unpublished).

Figure 2 shows a selection of the results obtained with patients 1 and 2. Using a mixture of radiolabelled probes I and IV, no specific bands were detected in *Eco*RI digests of DNA in either patient as compared with the normal control (Fig. 2A). Restriction enzyme digests of the DNA of patients 1 and 2 with four other enzymes (*Pst*I, *Bgl*II, *Hind*III and *Bcl*I) probed with probes II and III also failed to show any specific hybridization bands (data not shown). Further, no bands were observed when the DNA of patient 1 was probed with the cDNA probe V. In order to control for the presence of hybridizable DNA on the Southern blot in the tracks corresponding to patients 1 and 2, we have mixed probe V with an irrelevant autosomal probe for the human apolipoprotein A1 (apo A1; ref. 4). This internal control (Fig. 2B) shows that patient 1 has the normal apo A1 gene characterized by band c but lacks the major 5-kilobase band d characteristic of the normal factor IX gene. We have preliminary evidence to support the hypothesis that band d (Fig. 2) derives from a region of DNA 3' to the probe IV domain shown in Fig. 1; two faint satellite bands on either side of band d may derive from the two exon regions shown in Fig. 1. It appears that these minor bands are also lacking in patient 1. Figure 2C is a parallel analysis similar to that of Fig. 2B but performed with the DNA of patient 2. It shows the presence

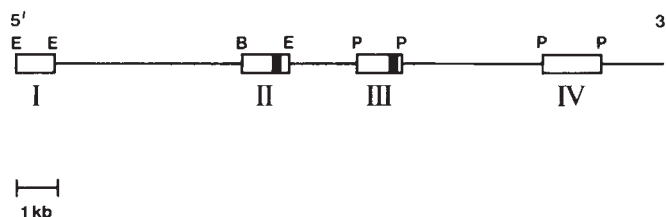


Fig. 1 Line diagram of part of the human genomic factor IX clone from which various specific probes were obtained. The original clone λ HIX1b contained repetitive elements, so unique probes were obtained by subcloning the regions I, II, III and IV which are free of such elements in the plasmid vector pAT153/*Pvu*II/8 (ref. 9). E, B, P refer to the restriction sites *Eco*RI, *Bgl*II and *Pvu*II. The filled-in sections in the rectangles marking probes II and III designate short exon regions. 5' and 3' show orientation with respect to mRNA.

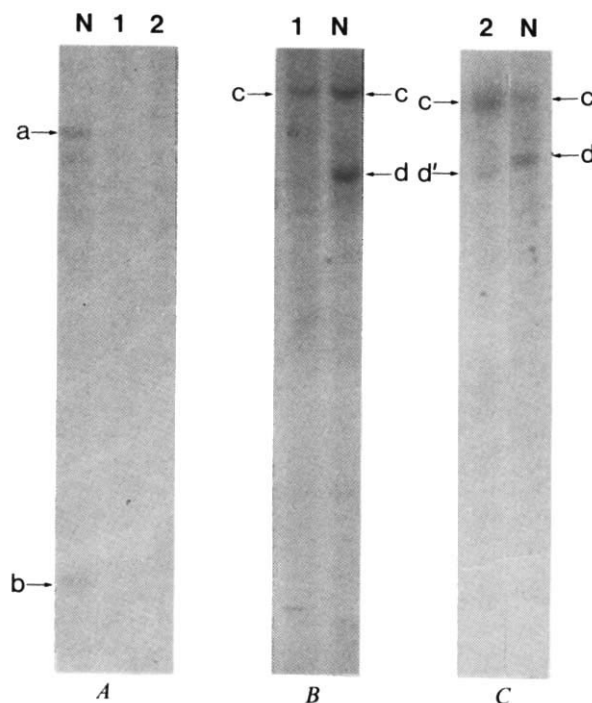


Fig. 2 Southern blots obtained from the DNA of patients 1 and 2, and normal controls (N). A, *Eco*RI digests hybridized with a mixture of probe I and IV. B, C, *Eco*RI digests hybridized with a mixture of probe V and the apo A1 probe, pBA1.2. a, b, 6.5- and 0.96-kilobase (kb) bands produced by hybridization of control DNA (track N) to probe IV and I, respectively. These are absent in patients 1 and 2. c, The 12-kb band produced by hybridization of the pBA1.2 probe to control and patients' DNA. d, d', 5- and 4.5-kb bands produced by hybridization of probe V to DNA of normal controls and patient 2, respectively. DNA was extracted from peripheral blood samples by established procedures⁸. The DNA was digested with the appropriate restriction enzyme electrophoresed in 0.8% agarose and transferred to nitrocellulose filters³. These were prehybridized overnight at 42 °C in 50% formamide, 5 × SSPE, 1 × Denhardt's solution and 0.1 mg of heat-denatured herring sperm DNA and then hybridized with nick-translated, heat-denatured probes ($\approx 0.5 \mu\text{Ci ml}^{-1}$) for 24–48 h⁸. The apo A1 specific probe was pBR322 containing a 1.2-kb segment of apo A1 cDNA⁴. Excess probe was washed off by incubation in 2 × SSC, 0.1% SDS for 20 min at 23 °C followed by 1 × SSC, 0.1% SDS and 0.2 × SSC, 0.1% SDS each for 1 h at 68 °C. Autoradiographs were developed after 3–7 days of exposure.

* Permanent address: The Paediatric Research Unit, Guy's Hospital, London SE1, UK.

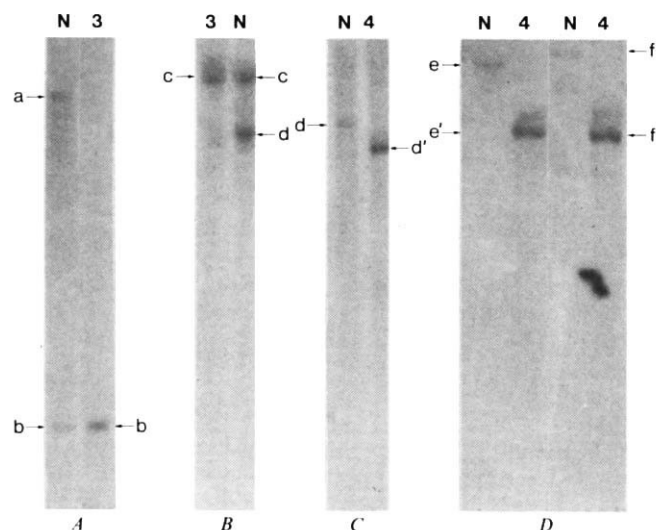


Fig. 3 Southern blots from the DNA of patients 3 and 4 and normal controls (N). *A*, *EcoRI* digests of patient 3 and control DNA hybridized to a mixture of probe I and IV. *B*, *EcoRI* digests of patient 3 and control DNA hybridized to a mixture of probe V and pBA1.2. *C*, *EcoRI* digests of patient 4 and control DNA hybridized to probe V. *D*, *PvuII* (first two tracks) and *HindIII* (last two tracks) digests of patient 4 and control DNA hybridized to probe V. *a*, *b*, 6.5- and 0.96-kb bands produced by probe IV and I respectively; *a* is absent in patient 3. *c*, *d*, 12- and 5-kb bands produced by probe pBA1.2 and V respectively; *d* is absent in patient 3. *d'*, 4.2-kb band produced by probe V in the track of patient 4. *e*, *e'*, 10- and 4.5-kb bands produced by probe V in the tracks of control and patient 4 respectively. *f*, *f'*, 15- and 4.4-kb bands produced by probe V in the tracks of control and patient 4 respectively.

of a weak band (*d'*) possibly of higher mobility than band *d* in the control.

Figure 3 shows Southern blots on patients 3 and 4. The DNA of patient 3, after digestion with *EcoRI*, showed normal hybridization bands with probes I and II but no bands with probes III, IV and V. Thus, in Fig. 3*A* and *B*, the patient's tracks show the bands expected from hybridization to probe I and the irrelevant apo A1 probe, but lack those due to probes IV and V. Results with probes II and III have been omitted for brevity. The DNA of patient 4 (a nephew of patient 3) shows bands hybridizing to all four genomic probes at the correct length but the size of the fragments hybridizing strongly to probe V is abnormal in *EcoRI*, *HindIII* and *PvuII* digests (Fig. 3*C*, *D*). Preliminary investigations of patient 5 with all five probes have shown no abnormalities.

Our results (summarized in Fig. 4) indicate that patients 1 and 2 have a deletion of more than 18 kilobases extending further than the region shown in Fig. 1. These are unlikely to be identical mutations because patient 2 contains a residual section of the gene, as a weak band appears when probe V is used. The patient 3 has a deletion greater than 9 kilobases but with part of the gene clearly present, and patient 4, despite being related to patient 3, has a different gross alteration in the region hybridizing strongly with probe V. The difference between the related patients 3 and 4 is surprising and will be the subject of a separate family study. At this stage, however, our incomplete knowledge of the normal gene structure limits our ability to make a more precise statement of the exact position or extent of the gene deletion. Indeed these deletions may be complicated by additions or inversions. However, no obvious rearrangements or deletions of the X chromosome could be seen in trypsin-Giemsa-banded⁵ metaphase preparations of lymphoblastoid cell lines established from patients 1, 2 and 4 (the other patients were not tested).

Various hypotheses could explain the origin of antibodies to factor IX in Christmas disease. These include lack of tolerance due to gross gene defects leading to absence of immunologically

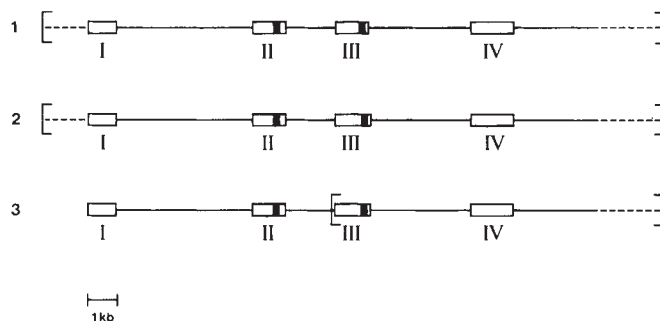


Fig. 4 Summary diagram of the extent of gene deletion in patients 1-3. The square brackets indicate the approximate location of the proposed deletions. The dashed lines (---) imply that the extent of the deletion in this region is uncertain.

recognizable factor IX, or breakdown of tolerance as a result of exposure to cross-reacting antigens, or the existence of factor IX allotypes. The first of these mechanisms seems the most important in the patients we have analysed since a gross gene deletion was present in at least 4 out of 5 patients with anti-factor IX antibodies. Gross gene deletion does not appear to be a general feature in 20 unrelated non-antibody-producing patients (our unpublished results). The primary risk of developing antibodies to factor IX would seem, therefore, to depend on the nature of the mutation and clustering among relatives of patients might be expected, although there is no clear evidence for this⁶ from family studies. But because of the small number of such antibody-producing patients, it is difficult to eliminate a familial predisposition.

The factor IX patients producing antibodies described here are the first showing gross structural changes of the factor IX gene. They provide a model for understanding the origin of the similar complication in classical haemophilia. This is the greater clinical problem because of the higher incidence of classical haemophilia and the fact that this complication develops in a higher proportion of cases with classical haemophilia than with Christmas disease⁷. We would predict gross changes of the factor VIII(C) gene in this group of haemophilia patients.

We thank the patients for donating blood samples, Drs P. Green, F. G. H. Hill, R. S. Mibashan and R. T. Wensley for allowing us to study their patients, Dr D. S. Anson for unpublished information, Dr F. E. Baralle for the apo A1 probe, and Professor J. H. Edwards for encouragement and advice. This work was supported in part by MRC project grants to G.G.B. and to Professor Edwards; F.G. was supported by the National Spastics Society and an anonymous donation; K.H.C. is a recipient of Uncle Bobs Travelling Scholarship and a C. J. Martin Research Fellowship from the National Health and MRC of Australia; D.J.G.R. was supported by a MRC research studentship.

Received 4 March; accepted 28 March 1983.

1. Choo, K. H., Gould, K. G., Rees, D. J. G. & Brownlee, G. G. *Nature* **299**, 178-180 (1982).
2. Rizza, C. R. & Spooner, R. J. D. *Br. med. J.* **286**, 929-933 (1983).
3. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
4. Shoulders, C. C. & Baralle, F. E. *Nucleic Acids Res.* **10**, 4873-4882 (1982).
5. Seabright, M. *Lancet* **ii**, 971-972 (1971).
6. Shapiro, S. S. & Holburn, R. in *Haemophilia and New Haemorrhagic States* (ed. Brinkhous, K. M.) (University of North Carolina Press, Chapel Hill, 1970).
7. Shapiro, S. S. in *Clinics in Haematology* Vol. 8 (ed. Rizza, C. R.) 207-214 (Saunders, London, 1979).
8. Maniatis, T., Fritsch, R. F. & Sambrook, J. (eds) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratories, New York, 1982).
9. Brownlee, G. G. & Choo, K. H. *British Patent Application* no. 8222486 (1982).

Erratum

In the article 'Three glycolytic enzymes are phosphorylated at tyrosine in cells transformed by Rous sarcoma virus' by J. A. Cooper, N. A. Reiss, R. J. Schwartz and T. Hunter, *Nature* **302**, 218-223 (1983), on page 221, the last line of the second column of text is incorrect. It should read: 'expected from the amino acid sequences'²⁶. In RSV-transformed ...

DNA and the doctor

Alan E.H. Emery

The New Genetics and Clinical Practice. By D.J. Weatherall.
Nuffield Provincial Hospitals Trust: 1983. Pp.133. £8.

IF ONLY all scientists could write like this! Professor Weatherall has produced a highly readable, straightforward account of a complex subject. He has avoided cluttering the text with a surfeit of references, no mean achievement when one considers that he is writing about a young discipline in which much is still not clear or is controversial.

The term "new genetics" refers to novel approaches in mapping and determining the fine structure of the human genome using recombinant DNA technology. The clinical application of this technology lies largely in the prevention of genetic disease through the detection of carriers and through prenatal diagnosis. In the more distant future, by isolating and cloning defective genes which are then induced to synthesize their products in the test tube, it will be possible to identify the basic biochemical abnormality concerned where this is not yet known. This approach, of so-called "reverse genetics", could eventually lead to the devising of effective treatments for such disorders. Professor Weatherall's main concern however is with understanding the fine structure of genes, particularly those involved in the haemoglobinopathies, and the avenues this knowledge opens up for prenatal diagnosis.

After an introductory chapter, Weatherall considers the frequency and clinical spectrum of genetic diseases. In round figures genetic disease and congenital malformations account for some 50 per cent of childhood deaths in hospital, and chronic diseases with a significant genetic component occur in about 10 per cent of the adult population. Thus genetic disease produces a considerable burden on the health services. Hopefully developments in recombinant DNA technology will, in various ways, help to reduce this burden in the next few years.

The real value of this little book lies in the masterly way in which the author describes, in relatively simple terms, a complex technology and the insights it has given us into molecular pathology. The single most important development which made the new technology possible was the use of restriction enzymes which cleave DNA at sequence specific sites. Essentially the new technology has allowed certain genes to be isolated, cloned in micro-organisms and their structure determined. In man we now know the detailed structure of, for example, the genes for the haemoglobins, immunoglobulins, insulin and growth hormone. Genes can no longer be considered simply as being discrete and

contiguous beads arranged on a string. They are preceded by critical regulatory sequences (RNA polymerase binding sites) and terminated by other sequences; further, almost all eukaryotic genes are interrupted by introns (intervening sequences) which are not transcribed into functional messenger RNA and protein. And between genes are large stretches of DNA with, as yet, no recognizable function — the so-called "selfish DNA".

By generating radioactive DNA sequences for use as probes, it is possible to demonstrate, by hybridization studies, the presence of homologous sequences in DNA extracted from tissues such as leucocytes and amniotic fluid cells. Information currently gained in this way can be applied to prenatal diagnosis in either of two ways. Directly, by demonstrating that a deletion is present, or that there is an altered restriction site within the abnormal gene; alternatively, and of more general application, indirectly, by showing that a mutant gene is in linkage disequilibrium (allele linkage) with a particular set of restriction sites — as in the case of sickle cell anaemia and β -thalassaemia in

Sardinia — or by demonstrating within families linkage with a restriction fragment length polymorphism. This latter approach is now a burgeoning area of research, especially with regard to disorders where the nature of the genetic defect is still unknown, such as Duchenne muscular dystrophy and Huntington's chorea.

The subject of gene therapy is dealt with only briefly, but Weatherall is quick to emphasize the problems and his sober account contrasts starkly with the sensationalism which has often accompanied discussions on this emotive subject. Finally, he considers the implications of the new genetics for clinical practice. Quite apart from considerations of cost-effectiveness and ethics associated with population screening and prenatal diagnosis, clinical practice in general is likely to change as the technology inevitably leads to fresh approaches to diagnosis, management and — ultimately — treatment. The medical profession as a whole will have to become informed about the technology, and clinical geneticists in particular will have to become very much involved in what is likely to revolutionize their approach to families with genetic disease over the next few years. For both groups, indeed for anyone with even a passing interest in genetics, Weatherall's book could not be a better starting point. □

Alan E.H. Emery is Professor of Human Genetics in the Medical School, University of Edinburgh.

Sorting out the sexes

Victoria A. Taylor

The Theory of Sex Allocation.

By Eric L. Charnov.

*Princeton University Press: 1982. Pp.355.
Hbk \$52, £34.50; pbk \$17, £11.30.*

It is one of the pitfalls of modern evolutionary thinking that whilst adaptive explanations spring readily to mind, they are all too often untestable. The study of sex ratios, however, proves a most fruitful testing ground for some of the predictions of evolutionary theory. The idea that the usual 1:1 sex ratio of males to females represents an evolutionary equilibrium resulting from the operation of natural selection is implicit in the work of R.A. Fisher. Sex ratio theory has since been elaborated and, more recently, some of its predictions tested, often using organisms whose sex ratios deviate from 1:1. For example, where local competition for mates occurs, a female-biased sex ratio is predicted, and this is indeed the case in many parasitoid wasps in which mating occurs on, or near, the host from which they have emerged.

But sex ratio is just one aspect of the more fundamental problem concerning the way in which resources are allocated to

male and female reproduction. How long, for instance, should a fish which changes sex during the course of its life, spend functioning as a male and how long as a female? How should an hermaphroditic plant divide its resources between production of pollen and ovules? Unlike sex ratio evolution, the broader question of sex allocation has been largely neglected until recently. Now, in *The Theory of Sex Allocation*, Eric Charnov draws on an extensive, though often anecdotal, literature to remedy this failing, exploring sex allocation from an evolutionary viewpoint through theory, observation and experiment.

Charnov's theoretical approach is based on finding the evolutionarily stable allocation between male and female reproductive functions. This simple basis is developed with a minimum of mathematical elaboration. It is then used to explore the conditions under which selection favours the evolution of the three main types of reproduction (dioecy, sequential hermaphroditism and simultaneous hermaphroditism) and the interesting evolutionary questions which each poses. For sequential hermaphroditism such questions concern the sex of an individual at first reproduction, whether it should then change sex and, if so, the age at which this change should occur. Theory predicts

protandry (male first) when, for example, female fecundity increases with size but males gain little by being larger. Protogyny (female first), on the other hand, is expected when only the largest males reproduce successfully because, for example, of intense inter-male competition. Such predictions, and more specific ones concerning the age of sex change, are strongly supported by quantitative data on pandalid shrimps and coral reef fishes (for a more detailed discussion, see *Nature* 301, 16; 1983). Relatively few data are available for other sex changers, such as marine molluscs and an annelid worm, but their inclusion, chiefly as potential tests for sex allocation theory, contributes to the general interest of the book.

Charnov draws on a yet wider assemblage of organisms in his analysis of the classic problem of sex ratio evolution in dioecious species and sex allocation in simultaneous hermaphrodites. This time it is mites and parasitoid wasps which provide the most convincing accord with theoretical predictions, whilst data from other groups, including nematodes, vertebrates, barnacles and higher plants, point to systems deserving further scrutiny.

The Theory of Sex Allocation is an important and stimulating contribution to modern evolutionary theory. In producing the first cohesive account of sex allocation, Charnov maintains a logical structure and lucid style which will make it accessible to the wide readership it deserves. □

Victoria Taylor is in the Department of Pure and Applied Biology, Imperial College, University of London.

Journals' review issue 1983

On October 6th *Nature* will publish the third review supplement devoted to science journals. The two previous supplements (covering journals first published between January 1978 and May 1980, and June 1980 and May 1981) appeared in *Nature* 293, 341-369 (1981) and 299, 491-514 (1982).

Criteria for inclusion of a journal in the 1983 issue are that:

- the first number appeared, or the journal was re-titled, between June 1981 and May 1982;
- it is published at least three times a year;
- the main language used is English.

Broadly, periodicals of professional interest to scientists will be considered for review, with the exception of abstracts' journals. Journals appearing prior to June 1981 but after January 1978, and not covered in the previous issues, will also be considered.

Publishers and societies are invited to submit four sample issues of journals satisfying the above criteria, including the first and most recent numbers, to the Review Editor, *Nature*, 4 Little Essex St, London WC2R 3LF, England (London 836 6633 ext 2570) as soon as possible.

The physicists

Rudolf Peierls

Haphazard Reality: Half a Century of Science.

By Hendrik B.G. Casimir.

Harper and Row: 1983. Pp.356. \$20, £12.50.

HENDRIK Casimir is one of the lucky generation of theoretical physicists who had their research training during the heady days of the start of quantum mechanics. While just too late to participate in laying the foundations of the new physics, he and his contemporaries, myself among them, learned the subject from the horses' mouths, from its originators. Casimir made many important contributions to the consolidation and the applications of quantum physics, of which the "Casimir operator", which has immortalized his name in the textbooks, is only one example.

In some ways he is an untypical theoretician. Theoretical physicists, particularly bright ones, are supposed to be complex personalities, prima donnas perhaps, or with excessive inferiority complexes, or some other mild madness. In their company Casimir stands out as remarkably sane. As might therefore be expected, his biography is undramatic, but what makes this book fascinating reading is the account of people, his teachers and colleagues and some others.

The portraits given here are very perceptive and lifelike, the author showing a remarkable memory for conversations and incidents that happened many years ago. His main teacher was Paul Ehrenfest, one of the most complex of physicists, but a brilliant teacher. Casimir writes with admiration and gratitude of Ehrenfest's inspiration and guidance, but does not gloss over his faults and occasional unreasonableness. He speaks with understanding of the motives that led to Ehrenfest's suicide, and he wonders whether he, as his young assistant, could have foreseen, and somehow prevented, this tragedy.

There is an excellent study of Niels Bohr, and here, too, the great admiration for the man which Casimir shares with practically all theoretical physicists who ever set foot in Copenhagen, does not prevent him from seeing the funny side, and characterizing Bohr's unique use of language. For this he is well qualified, as the author of a famous essay on "Broken English" as the language of communication in science. Then there is Pauli. Casimir was his assistant for a year, and survived the experience in spite of Pauli's merciless, critical remarks for which he was famous. But Pauli, too, comes out likeable, if often comic.

Of Casimir's own generation, there is above all Landau, with an impressive command of physics, and apt to make earnest theories about life, politics and physicists, as well as about physics. There is Gamow,

always taking any opportunity of making a joke, but with a good judgement on unsolved problems. And there is Houtermans, eternal bohemian, married four times, but with only three wives.

The great men of Dutch physics are also included — Kamerlingh Onnes, the founder of the Leiden laboratory, and the first to liquefy helium, who is shown to have feet of clay; de Haas, erratic and unpredictable; Keesom, with a huge body, but a nimble mind and a surprisingly great command of physics; and Gorter, a great experimentalist, perhaps undervalued by many, including himself.

In the 1930s, Casimir toyed at one time with the idea of accepting a job in Hitler's Germany, and turned it down only after a

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Hendrik Casimir — stands out among theoretical physicists as remarkably sane.

message from Niels Bohr. This episode is recounted with great honesty and self-questioning, as is his conduct during the War, when he remained in the Netherlands but was fortunate in that he and his family escaped serious hardship. He worries whether he was to blame for not joining the resistance or otherwise fighting the occupation, but then he felt responsible for his family. During the War he was offered a job in the research department of Philips, which he accepted mainly because the German occupation and personal disagreements had made his work in Leiden unattractive. But he enjoyed the work at Philips, and stayed on, to become joint director of the laboratory, and to join the board of the company.

Casimir talks with great frankness about his attitude to industry, realizing that the life of an industrial tycoon was what as a young man he would have looked down upon as unworthy. He gives a short history of the firm, and describes some of the people he met, but refrains from going into details of his work, for obvious reasons. Two final chapters give his views on the development of particular lines in industry which have depended strongly on recent

progress in science, and on the relation between science and technology in general.

In outlining his academic career, and his contacts with his teachers and colleagues, Casimir has to refer to the state of physics during that period, and to provide the background he includes a chapter on the development of physics. This is a very brief summary, naming the salient steps rather than trying to explain them; this approach can be defended since the reader cannot expect to be taught the whole of modern physics incidentally. Three appendices, however, give more detail for the physicist reader.

The introduction says "... I have tried to write about people I met and knowledge I gained rather than about myself".

Casimir faithfully carries out this intention and keeps himself in the background. But on finishing the book, we feel we have got to know and like him. His life is undramatic, as I have said, but we see his views and his principles emerge from the narrative and, occasionally, his rather puckish sense of humour. And we know what it felt like to be a physicist in the exciting days of the beginning of quantum mechanics.

The title *Haphazard Reality* evidently relates to the statistical nature of modern physics; it is, perhaps, also meant to hint at important chance influences in Casimir's own life. □

Sir Rudolf Peierls is Emeritus Professor of Theoretical Physics at the University of Oxford.

Breadth and depth of plant carbohydrates

P.H. Rubery

Encyclopedia of Plant Physiology, Vols 13A and 13B. Vol. 13A Intracellular Carbohydrates, Vol. 13B Extracellular Carbohydrates.

Edited by F.A. Loewus and W. Tanner. Springer-Verlag: 1982. Vol. 13A, pp.918, DM298, \$132. Vol. 13B, pp.769, DM268, \$125.

THESE two, large, multi-author volumes (1,700 pages in all) are divided for convenience by the plasma membrane. They deal, from monomer to polymer, with the biochemistry and selected physiological aspects of, respectively, intra- and extracellular carbohydrates (interpreted to include glycoconjugates and hydrophobic wall components). Coverage of carbohydrate metabolism as such is mostly concerned with aspects of biosynthesis and early committed steps in degradation; thus glycolysis and photosynthetic carbon pathways are to be found in other parts of the *Encyclopedia*. Also, it is disappointing to find little explicit treatment of recent methodological developments in the structural analysis of polysaccharides.

Volume 13B is divided into sections on the cell walls of higher plants (270 pp.) and of algae and fungi (200 pp.), carbohydrate secretion (100 pp.), and surface recognition phenomena (100 pp.). Each volume has its own author and subject index which, however, lack cross-references between the two books.

As one would expect, the well-researched core topics such as sucrose, starch, glycoproteins (figuring in both volumes) and cell walls are thoroughly covered in authoritative reviews similar to those found in other recent encyclopaedic publications such as the *Comprehensive Treatise on Plant Biochemistry* (reviewed in *Nature* 299, 762; 1982). Although many

of these chapters retain a freshness of approach — evident, for example, in the intelligent, analytical and sceptical account of starch storage (Chapter 20 in Vol. 13A) — a degree of dullness and overkill has crept in. Thus the first four chapters of Vol. 13B, on macromolecular structure and biophysical aspects of higher plant cell walls, are little more than short, overlapping summaries and would have benefited both from some expansion and from the editorial wielding of Occam's Razor.

For me, much of the value of publications such as these lies in the cross-fertilization which can result when updated archival material and less-familiar and less-developed topics, such as algal autolysins (Vol. 13B), are brought together within the same covers. Collation of scattered data and their integration in a contemporary context is one function which is well performed in the accounts of glycoproteins, steryl glycosides and non-starch reserve polysaccharides. There are also numerous contributions, such as those dealing with sugar alcohols and nectar secretion, which reveal the importance and fascination of areas sometimes (unjustly) regarded as obscure and prosaic. The chapters dealing with carbohydrate transport (Vol. 13A) and with cell surface phenomena (Vol. 13B), including plant-pathogen interactions, are particularly successful and illustrate the importance of including functional aspects to maintain the balance of a work of this sort.

Perhaps when the third series of the *Encyclopedia* is published sometime in the next millennium, the mechanism of cellulose microfibril deposition and other celebrated puzzles — which are brought out explicitly or, unfortunately more often, implicitly by the current authors — will be a little nearer elucidation. In the meantime, these books will be generally invaluable sources of information and inspiration for research work and advanced teaching. □

P.H. Rubery is a Lecturer in Biochemistry at the University of Cambridge.

Angles on creativity

Morris Kline

Geometrical Investigations Illustrating the Art of Discovery in the Mathematical Field.

By John Pottage.

Addison-Wesley: 1983. Pp.480. \$35.95, £26.95.

THE ART of discovering mathematical theorems, to say nothing about proof, has caught the interest of several authors. Some of the resulting books have treated the art in a discursive form and so have failed to pin down the helpful steps such as tackling simple cases first, generalizing on simpler cases, testing hypotheses through comparison with related situations — for example using polygons to suggest a result desired for a circle, which loosely described is an infinite-sided polygon — the use of analogy and other techniques. A few books, notably those of George Polya, have been more specific in their suggestions.

Creativity in mathematics can rarely be the outcome of a chance discovery or plain hard work as Gauss and Einstein very modestly declared. Creation involves groping, blundering, conjecturing, chance associations, experimentation and hypothesizing. It also calls for mental qualities of imagination, insight, intuition, divination, curiosity, hard work, patience, luck and even an aesthetic sense. Asking the right question is another feature of the process; originality, depth and breadth are marks of worth. It is not divine revelation, though the creative mind may at times seek to see itself enshrined as such after the act of creation.

John Pottage is realistic. Though a few Einsteins may have possessed some genius, almost all mathematicians must and do use the down-to-earth methods of creation or discovery treated and exemplified by Pottage. He is specific in his examples of discovery and his primary aim is to inform the layman and to teach teachers how to help students discover. He rightly opposes the rule-of-thumb procedures for learning mathematics or facing students with a correct theorem and a deductive proof as a *fait accompli*. Though he starts with a difficult problem — namely, to find the formula for the perimeter of an ellipse — he considers many simpler geometric problems involving the perimeters, areas and volumes of polygons, polyhedra, circles, spheres, cylinders and cones; by using — as opposed to discussing — some of the techniques mentioned above, he shows how discovery of the correct hypothesis can be reached. Many of his examples involve curvilinear figures. These illustrations of discovery are not simple, yet the questions he raises and the modifications he suggests in order to reach the correct theorem are to the point and clever. No general principles of

discovery are explicitly formulated in the book but must be inferred from the techniques used to arrive at such principles. Nonetheless the examples Pottage has chosen are certainly illustrative of how one might proceed.

Pottage takes over from Galileo the dialogue form and, indeed, uses the very same names for his characters that Galileo used. The wisest character, Salviati, takes up ideas put forward by his more naïve fellows such as Simplicio, and by incisive questions and suggestions gradually leads the several characters to the correct hypotheses. Though this dialogue form makes for interesting reading, it is at times too verbose for my taste and on occasion distracts from the principal objectives of the book. This approach, however, certainly avoids the rather dry, didactic presentations to be

found in pedagogical texts. Some of the examples do involve rather advanced algebra and even some calculus. Hence the book includes much more than might be useful to secondary school teachers. But college level teachers will certainly be able to follow this portion of the contents.

There are many excellent historical notes in an extensive appendix and historians of mathematics will certainly find these helpful. As a final bonus the author offers in another appendix a number of novel geometric problems, with hints and solutions, which a student might use to practise discovery. □

Morris Kline is Professor Emeritus of Mathematics at the Courant Institute of Mathematical Sciences of New York University. His most recent book is Mathematics, the Loss of Certainty (Oxford University Press, 1980).

Men of the mountains

Roy Porter

Geology in the Nineteenth Century: Changing Views of a Changing World.
By Mott T. Greene.
Cornell University Press: 1983. Pp.324.
\$29.50, £23.50.

WE HAVE been witnessing a long-overdue revival of interest in the history of geology, especially in its heroic age between Hutton and Darwin. The old-fashioned "great man" approach (e.g. Geikie's *Founders of Geology*) has yielded to the study of problems (e.g. Rudwick's *The Meaning of Fossils*). Yet shortcomings have persisted. Many eras have remained unilluminated, not least the problem of what happened "after Darwin"; and research has remained Anglocentric and parochial. Mott Greene's powerful and original book, however, goes a long way towards putting that right. His is the first interpretation of the science's maturation which is genuinely international, which goes right up to the First World War and which refreshingly gets away from old warhorses such as the triumph of Uniformitarianism.

As Greene sees it, Hutton, Werner and their pupils equipped geologists with a tool-kit of terms and methods, but they failed to solve the key problems, especially the one which — particularly in Romanticism's heyday — was staring them in the face: mountains. How to explain not just individual peaks, but that fact that continents were ribbed by chains of them. Werner's aqueous theory of sedimentation had made more sense of stratigraphical order than of deformation; and for its part, Lyell's advocacy of gradual, actual causes and his stress on denudation were to explain mountain relief better than mountain structure.

Greene argues that two basic solutions for orogenesis were offered in the nineteenth century. One tradition, building on

Lyell, subsumed it within fundamental equilibrating cycles of erosion and sedimentation, producing secular oscillations — subsidence and elevation — in the distribution of land and sea. From John Herschel and Babbage in the 1830s, via the American James Hall, up to the "isostasy" of Dutton and Gilbert, mountain-building was seen by this school as regular and piecemeal, an epiphenomenon of the crust's basic homeostatic adjustment.

But explaining continents wasn't explaining mountains, claimed the other camp. Actualism and isostasy might account for the rise of Scandinavia out of the Baltic, but not of the Alps. Vast in their ranges, commanding in scale and bafflingly complex in their structural deformations, mountain chains were evidence, rather, of paroxysmal episodes. In the 1820s, deluges were replaced as the favourite cataclysmic agent by the global secular cooling advocated by La Place and Cordier, which allegedly produced contraction. Elaborated most grandly by Elie de Beaumont, this hypothesis was fruitfully applied by the Rogers brothers and James Dana in their studies of the Appalachians. Eventually it was to suffer its own collapse, internally the victim of its own grandiosity (Elie de Beaumont claimed that mountain systems followed a global pentagonal geometry), and externally killed by geophysics. For, especially after Joly's revelation of internal radioactivity, geophysics could no longer accept that crustal contraction was of a scale adequate to explain crustal deformation.

While both of these camps were making "the poor old earth play tricks", Eduard Suess was toiling away in Vienna at his noble synthesis, stretching from *Die Entstehung der Alpen* in 1875 to the majestic *Das Antlitz der Erde* which appeared in four volumes between 1883 and 1904. Striving to bridge these rival traditions while avoiding oversimplification, Suess brought forth order out of the diversity of crustal movements, dignifying tectonics in the process as the study of structural dynamics,

and treating orogenesis under general tectonics. Regarding mountain-building forces as transcontinental lateral thrusts, Suess pictured the Alps, along with the Carpathians and Asiatic chains, as part of a gigantic north-moving fold.

Suess's solution echoed, of course, the researches of field-workers; facing geology's riddle of the Sphinx, they had been staggered by the vast Alpine structural deformations and metamorphoses. Some puzzles — overthrusting of older onto younger sediments, underfolding and the massive thrusting of recumbent anticlinal ridges — found solution in the *nappe* theory of Bertrand, Termier, Schardt and Lugeon. *Nappe* is French for table-cloth. Push one across a flat surface; not only does it fold, but the later folds rise up over the earlier ones; likewise with Alpine strata, which, however, are additionally modified by fracture and erosion. For Bertrand, the Alps did not just contain overthrusts; they *were* overthrusts.

The *nappe* theory made good field sense. But, as Greene shows, it lacked a causal theory. Here, a new burst of theoretical energy around the turn of the century promised to come to the rescue, solving the riddle of vast lateral pressures. Abandoning crustal contraction, it postulated the fracture and movement of great palaeocontinents, leading to ideas of continental drift in the work of Haug, Frank Taylor and of course Wegener (not the least strength of Greene's book is that it puts paid to Wegener as a lonely genius, by rooting his ideas directly in contemporary tectonics). Yet these concepts remained inchoate and neglected. Geology entered the present century racked by theoretical pluralism and confusion, lacking — despite Suess — an overarching theory.

There was much more to "geology in the nineteenth century" than Greene covers: stratigraphy, mapping, palaeontology and so on are hardly mentioned. And even his account of mountain-building is selective, saying little about the roles of geochemistry and petrology, for instance. But as a history of tectonic thinking this is a major work — sure-footed in handling complexities, astute in its attention to both theory and practice, ideas and individuals; international in scope; proficient in its geology, yet alert to the impact of wider intellectual currents and other sciences. As a historian, Greene steers confidently between the Scylla of Whiggism and the Charybdis of modish philosophies of science, and knows how to shape a robust narrative. There are, of course, points to criticize. Some important recent publications are ignored, and the text is sadly littered with errors in proper names, titles and so on. Yet this book will deservedly take its place as a milestone in the history of geology. □

Roy Porter is a Senior Lecturer at the Wellcome Institute for the History of Medicine, London. Among other books he is author of The Making of Geology (Cambridge University Press, 1977).